YALE ECONOMIC ESSAYS

VOLUME 2 NUMBER 2

Determinants of Productivity Change in United States Manufacturing — *Benton F. Massell*

Connections Between United States' and Canadian Capital Markets, 1952–1960 — *Gerald K. Helleiner*

Patterns of Growth and Valuations in a Two-Sector Model — *Mordecai Kurz*

Inflationary Financing and Economic Development: The Pakistan Experience, 1951–1959 — *Parvez Hasan*

FALL 1962

YALE ECONOMIC ESSAYS

This journal is intended to serve two principal functions. All doctoral dissertations in economics at Yale University will be published in its pages either in full or in extended summary. It is also a medium for economic comment and analysis by faculty, students, and graduates in economics of Yale University.

Editorial Board: Neil W. Chamberlain
Raymond P. Powell Arthur M. Okun

Editorial Staff: Bruce W. Morgan and Karen H. Hester

Volume 2 Fall 1962 Number 2

Yale Economic Essays is printed at the Carl Purington Rollins Printing-Office of the Yale University Press. A one-year subscription (two issues) is $4.00. Single copies are $2.50. The rate for students for one year is $2.00. Checks should be made payable to Yale University. Please address any correspondence to the Editorial Board, 37 Hillhouse Avenue, New Haven, Connecticut.

Copyright 1962 by the Yale Economic Essays.

YALE ECONOMIC ESSAYS

TABLE OF CONTENTS

Determinants of Productivity Change in United States Manufacturing — *Benton F. Massell* — 303

Connections Between United States' and Canadian Capital Markets, 1952–1960 — *Gerald K. Helleiner* — 351

Patterns of Growth and Valuations in a Two-Sector Model — *Mordecai Kurz* — 403

Inflationary Financing and Economic Development: The Pakistan Experience, 1951–1959 — *Parvez Hasan* — 475

Summary: On the basis of a conceptual distinction between capital deepening and technical change, the latter is found to be the dominant factor in accounting for increases in output per manhour in U.S. manufacturing during the past four decades. Moreover, the rate of technical change in aggregate manufacturing is found to have exceeded a weighted average of the industry rates of technical change by a substantial margin. The discrepancy between the two, which has been termed "interindustry technical change," can be attributed largely to shifts of capital from industries where the marginal productivity of capital is low to those where this marginal productivity is higher. Finally, it is seen that industries with high technology levels tend to have both high capital turnover ratios and high wage rates, lending support to the hypothesis that investment acts as a vehicle for bringing about technical change.

This essay was submitted to Yale University in 1962 in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The author is currently on the research staff of The RAND Corporation.

Dissertation Committee: Professors William J. Fellner (Chairman), Arthur M. Okun, and Michael C. Lovell.

*

This study was begun while I was working with the Cowles Foundation for Research in Economics; it was subsequently continued under a Ford Foundation Fellowship, and was completed at The RAND Corporation. Computations were financed by the Cowles Foundation and The RAND Corporation. I am grateful to the members of my dissertation committee and to many colleagues at Yale and at RAND for useful suggestions. I am particularly indebted to Arthur Okun, Alan Manne, Richard R. Nelson, and Sidney G. Winter, Jr.

B.F.M.

DETERMINANTS OF PRODUCTIVITY CHANGE IN UNITED STATES MANUFACTURING

Benton F. Massell

1. Introduction

The past forty years have witnessed a significant rise in output per manhour in all major segments of the U.S. economy. The present study attempts to contribute to our understanding of the forces underlying this trend.

It is convenient to begin by classifying these forces as either (1) capital deepening (a rise in the ratio of capital to labor employed in production, where capital is defined as all non-human productive inputs) or (2) technical change (any change in the technique of production which enables a larger output from a given combination of inputs). Conceptually, the distinction is quite clear between technical change, on the one hand, and an increase in the capital-labor ratio, on the other. Expressing output per manhour as a function of capital per manhour, increased capital intensity corresponds to a movement along the production function, while technical change refers to a shift from one function to another. Statistically, the distinction is obscured by the fact that, while a production function represents a range of hypothetical alternative factor combinations, in fact, at any one time, only one combination can be observed. If the function is shifting over time due to improvements in the technology, only one point on each function is observable, and the effects of technical progress and increased capital intensity on output per manhour are confounded.

Section II of this essay presents a method of apportioning past changes in productivity (i.e., the average productivity of labor) between technical change and capital deepening. This procedure is applied to the manufacturing sector of the U.S. economy for the period 1919 to 1958.

In Section III, the same analysis, with appropriate modifications, is applied to the Standard Industrial Classification (SIC) two-digit manufacturing industries for the postwar period, and the industry rates of technical change are compared with aggregate figures to determine the relative importance of "interindustry" and "intraindustry" technical change. The former is shown to result from shifts of capital and labor among industries.

The model discussed in Section II was first presented formally by Robert Solow[1] in estimating the quantitative importance of technical change in the private non-farm segment of the United States economy. Solow attributed 90 per cent of the rise in productivity to improvements in the technology, and the remaining 10 per cent to capital deepening.[2] The results

1. Robert M. Solow, "Technical Change and the Aggregate Production Function," The Review of Economics and Statistics, XXXIX (August 1957), pp. 312-20.

2. In fact, Solow's initial figure was 87.5 per cent, but computational errors were subsequently found in his work. See W. P. Hogan, "Technical Progress and Production Functions," ibid., XL (November 1958), pp. 407-08.

of his and similar studies[1] supported the view that technical advance was of overwhelming importance in bringing about the observed long-run increase in output per manhour.[2]

However, these studies fail to consider the possible interaction or complementarity between shifts in the production function and investment. Section IV of the present essay examines this question of complementarity in some detail. A model is developed which explicitly recognizes that technical change may depend on gross investment. The hypotheses suggested by this model are tested empirically, using time series and cross section data.

The final section of the essay summarizes the conclusions which are supported by the empirical results of the earlier sections and briefly explores some of the implications of these conclusions.

2. Technical Change vs. Capital Deepening[3]

This section presents estimates of the relative importance of technical change and capital deepening in bringing about productivity increases in United States manufacturing.

The Model

Assuming technical change to be neutral (i.e., that it shifts the production function iso-elastically), the production function can be written:

$$(2.1) \quad Q = A(t)f(K, L),$$

where Q = output, L = labor input (measured in manhours), K = capital input, t = time, and A can be considered as an index of the level of

1. See, for example, Solomon Fabricant, "Economic Progress and Economic Change," 34th Annual Report of the National Bureau of Economic Research (1954); M. C. Urquhart, "Capital Accumulation, Technological Change, and Economic Growth," Canadian Journal of Economics and Political Science, XXV (November 1959), pp. 411-30; M. Abramovitz, "Resources and Output Trends in the U.S. since 1870," American Economic Review, Papers and Proceedings, 46 (May 1956), pp. 5-23; J. Kendrick, "Productivity Trends: Capital and Labor," Review of Economics and Statistics, 39 (August 1956), pp. 248-57; O. Niitamo, "Development of Productivity in Finnish Industry, 1925-1952," Productivity Measurement Review, 15 (November 1958), pp. 30-41; J. Schmookler, "The Changing Efficiency of the American Economy, 1869-1938," Review of Economics and Statistics, 34 (August 1952), pp. 214-321; and T. W. Schultz, "Reflections on Agricultural Production, Output and Supply," Journal of Farm Economics, 38 (August 1956), pp. 748-62.

2. Regarding the danger of attributing causal significance to these results, see below, pp. 346-47.

3. This section is a revision of a paper published in the Review of Economics and Statistics, XLII (May 1960), pp. 182-88.

technology. The neutrality assumption permits writing A(t) as a factor of proportionality (instead of having to express it as an argument in the function). If it is further assumed that the function is linear homogeneous, one can write:

$$(2.2) \quad q = A(t)f(k, 1),$$

with $q = Q/L$ and $k = K/L$. Solow has shown[1] that by assuming competitive pricing in factor markets, an expression for the percentage rate of technical progress is:

$$(2.3) \quad \frac{\dot{A}}{A} = \frac{\dot{q}}{q} - \beta \frac{\dot{k}}{k}$$

where β = property's share of the product, and the dots denote derivatives with respect to time.[2] If discrete changes are substituted for time derivatives, (2.3) can be rewritten:

$$(2.3a) \quad \frac{\Delta A}{A} = \frac{\Delta q}{q} - \beta \frac{\Delta k}{k},$$

---

1. Solow, pp. 312–13.

2. The derivation of (2.3) is as follows: Taking logarithmic derivatives of (1) gives:

$$(2.2a) \quad \frac{\dot{Q}}{Q} = \frac{\dot{A}}{A} + \frac{1}{f}\left\{\frac{\partial f}{\partial K} \cdot \dot{K} + \frac{\partial f}{\partial L} \cdot \dot{L}\right\}.$$

Noting that $f = \frac{Q}{A}$, this can be rewritten:

$$(2.2b) \quad \frac{\dot{Q}}{Q} = \frac{\dot{A}}{A} + A\left\{\frac{K}{Q} \cdot \frac{\partial f}{\partial K} \cdot \frac{\dot{K}}{K} + \frac{L}{Q} \cdot \frac{\partial f}{\partial L} \cdot \frac{\dot{L}}{L}\right\}.$$

Competitive factor pricing enables writing

$$(2.2c) \quad \beta = A \cdot \frac{K}{Q} \cdot \frac{\partial f}{\partial K},$$

while linear homogeneity ensures that

$$(2.2d) \quad 1 - \beta = A \cdot \frac{L}{Q} \cdot \frac{\partial f}{\partial L}.$$

Thus, substituting (2.2c) and (2.2d) into (2.2b),

$$(2.2e) \quad \frac{\dot{Q}}{Q} = \frac{\dot{A}}{A} + \beta\frac{\dot{K}}{K} + (1 - \beta)\frac{\dot{L}}{L},$$

from which (2.3) follows easily. Solow pointed out (op. cit., p. 313) that the exhaustive classification of factors, together with the assumption of competitive factor pricing, implies linear homogeneity as well.

and an index of technology for each year can be derived from the identity

$$(2.4) \quad A(t+1) = A(t)\left[1 + \frac{\Delta A(t)}{A(t)}\right],$$

setting $A(t_1) = 1$. This index will show the position of the production function at each point in time.

The assumption that factors are competitively priced may be questioned, particularly during either war or depression. Nevertheless, the marginal productivity doctrine does not have to be taken too literally as an instantaneous equilibrating mechanism; rather, this doctrine can refer to a tendency for factors, in the long run, to be competitively priced, so that in most years the rate of return to the factor will serve as a first approximation to its marginal product. When property's share is higher than is warranted by its marginal product, due to short-run disequilibrium in factor markets, then technical change will be understated or overstated accordingly as the percentage rate of increase in capital is greater or less than zero. The results of this study are not greatly affected by small changes in property's share.

There is little evidence either to support or to reject the neutrality assumption. The relative constancy of labor's share of net income has been advanced as evidence that technical progress has been, on the whole, neutral,[1] but this condition is neither necessary nor sufficient unless the functional form of (2.1) is further specified.[2] One must either postulate a specific type of production function or assume neutrality; we have chosen the latter because it appears less restrictive.

The linear homogeneity assumption should not be particularly objectionable; it is one commonly found in economic analysis. Land is not of great importance in manufacturing, so one should not expect to find significantly diminishing returns to labor and capital. One is more likely to be confronted with increasing returns, and if this is the case, the benefits of increases in scale will be attributed to technical advance.

The Variables

To compute the technology index, observations are needed on four variables: output, capital, labor, and property's share. As none of these is wholly unambiguous, it might be useful to discuss at some length precisely what is called for in the model.

One interpretation of equation (2.1) is that it is analogous to an engineering relationship, with both output and inputs referring to physical

1. This is essentially the meaning of Solow's test for neutrality (op. cit., p. 313).

2. See S. G. Winter, Jr., "Testing for Neutrality of Technical Change," Cowles Foundation Discussion Paper No. 61 (multilithographed; New Haven: Cowles Foundation for Research in Economics at Yale University, 1958).

goods. A gross concept of output would be appropriate to this scheme, and property's share would include depreciation allowances. However, if (2.1) is viewed in this way, it becomes necessary to include a third factor of production, viz., raw materials, or more generally, purchases from outside of the sector; this, however, violates the assumption of constant returns to labor and capital. By including outside purchases as an input, labor's share plus capital's share fail to exhaust the product.

Alternatively, (2.1) can be considered as a value relationship. Here two choices present themselves: Output can be considered as net income originating in the sector (i.e., the sum of net factor incomes), or as value-added (including deprecation). The first of these is conceptually more appealing, for if one is concerned with output in value terms (rather than physical units), then it makes little sense to include as an increase in output that which results from increased durability of capital. One is, unfortunately, limited by the type of data available, particularly due to the treatment of capital consumption. Depreciation, which is deducted from gross output to obtain net output, is, to a large extent, arbitrary, depending on the conventions pursued by accountants and on prevailing tax laws. Net output would be satisfactory only if it measured the sum of factor incomes based on a notion of capital consumption which expressed actual deterioration of the services associated with structures and buildings. Available data are quite inadequate in this respect.

While in a "normal" year, i.e., a peacetime, full-employment year, depreciation charges may give a fair approximation to capital consumption, in a "troubled" year such as 1930, when output has fallen considerably below the expected, or normal, value, depreciation in the accounting sense may greatly exceed the actual decline in the value (in real terms) of the capital stock.[1] While in some cases depreciation is a function of time only, in other cases it is a function of both time and output. While a straight line depreciation allowance may serve as a good approximation to actual capital consumption when actual output equals expected output, it is a poorer approximation when output either greatly exceeds or falls short of its expected value.

It appears preferable, consequently, to use some other measure of output as a proxy for net output. As a carefully worked series relating to gross physical output per manhour has recently been completed by John Kendrick, the natural choice is to use this source.[2] This series appears in Column 1 of Table 1 and in Figure 1. The distinction made by Kendrick between physical output and valued output is that the former represents an aggregation over physical goods produced, weighted by their respective

1. During the 1930's, straight line depreciation was predominantly used.

2. J. Kendrick, Productivity Trends in the United States (Princeton: Princeton University Press, 1961).

TABLE 1

Output per Manhour (q), Capital per Manhour (k), and Labor's Share of Income Originating (β) in U.S. Manufacturing, 1919–1958

Year	Output per Manhour (Index Numbers, 1929 = 100.0)	Capital per Manhour (Index Numbers, 1929 = 100.0)		Labor's Share (Per cent)
		Unadjusted for Idle Capacity	Adjusted for Idle Capacity	
1919	58.0	66.1	66.8	75.47
1920	61.5	70.3	69.7	73.30
1921	71.0	104.9	95.4	79.17
1922	80.4	94.9	90.6	80.11
1923	77.4	84.6	84.6	77.94
1924	82.3	97.3	95.0	79.42
1925	87.7	95.0	94.2	76.59
1926	89.4	95.3	96.6	74.26
1927	91.5	99.0	98.1	78.28
1928	95.6	102.1	100.8	77.10
1929	100.0	100.0	100.0	75.12
1930	103.7	122.0	115.1	77.64
1931	103.9	148.5	129.0	89.01
1932	97.1	176.8	139.5	109.09
1933	105.7	155.9	121.0	105.75
1934	110.4	143.8	116.3	90.19
1935	117.6	126.6	104.5	83.36
1936	118.5	110.0	94.5	79.01
1937	116.9	106.1	93.9	79.37
1938	114.9	135.1	113.0	84.14
1939	125.8	115.1	98.5	80.84
1940	131.9	106.7	94.2	74.26
1941	136.7	88.5	82.4	69.68
1942	139.1	76.1	74.9	71.80
1943	141.0	64.6	65.5	74.01
1944	139.4	64.9	66.3	75.52
1945	137.5	75.3	76.3	78.14
1946	126.4	89.4	88.8	79.77
1947	133.8	92.0	91.7	76.80
1948	138.6	96.8	96.6	73.66
1949	143.7	108.9	106.4	74.52
1950	155.3	103.4	101.5	71.56
1951	153.3	101.9	102.1	71.41
1952	157.1	106.4	107.0	75.67
1953	163.9	106.0	116.0	77.27
1954	168.5	118.1	115.9	79.01
1955	179.7	113.5	112.6	75.46
1956	183.2	116.6	115.9	77.80
1957	187.1	123.9	122.5	78.84
1958	193.3	136.8	131.7	81.57

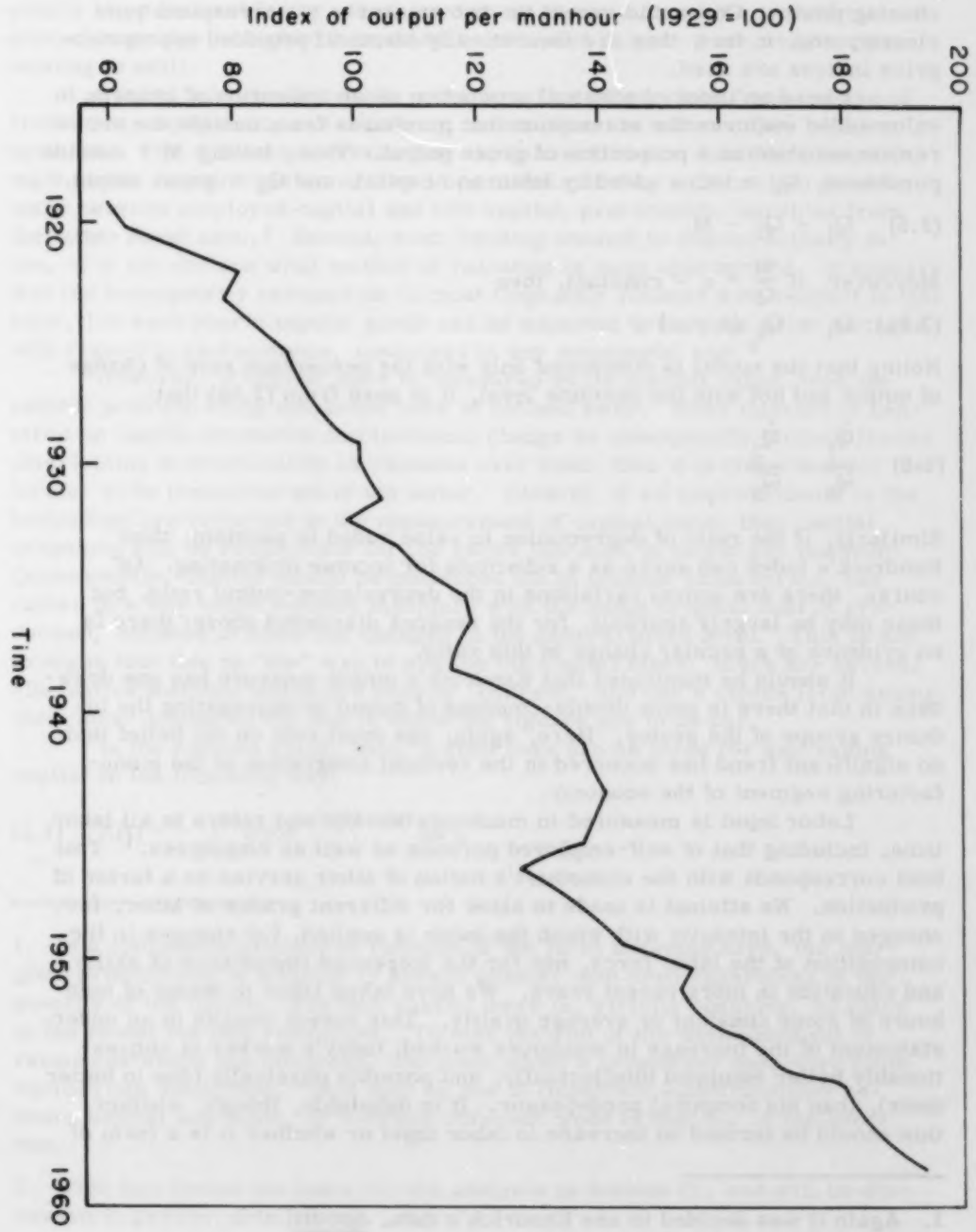

Fig. 1—Output per manhour in U.S. manufacturing, 1919-1958

market prices in a base year, and expressed as an index number, while the latter is a deflated value series, i.e., one reduced to dollars of constant purchasing power. One would expect the two measures to correspond quite closely, and, in fact, they are theoretically identical provided appropriate price indices are used.

Use of an index of physical production as an indication of changes in value added requires the assumption that purchases from outside the sector remain constant as a proportion of gross output. Thus, letting M = outside purchases, Q_1 = value added by labor and capital, and Q_2 = gross output,

(2.5) $Q_1 = Q_2 - M.$

Moreover, if $\frac{M}{Q} = c$ = constant, then

(2.5a) $Q_1 = Q_2 (1 - c)$.

Noting that the model is concerned only with the percentage rate of change of output and not with the absolute level, it is seen from (2.5a) that:

(2.6) $$\frac{\dot{Q}_1}{Q_2} = \frac{\dot{Q}_2}{Q_2} .$$

Similarly, if the ratio of depreciation to value added is constant, then Kendrick's index can serve as a substitute for income originating. Of course, there are annual variations in the depreciation-output ratio, but these may be largely spurious, for the reasons discussed above; there is no evidence of a secular change in this ratio.

It should be mentioned that Kendrick's output measure has one drawback in that there is some double-counting of output in aggregating the industry groups of the sector. Here, again, one must rely on the belief that no significant trend has occurred in the vertical integration of the manufacturing segment of the economy.

Labor input is measured in manhours worked and refers to all labor time, including that of self-employed persons as well as employees.[1] This best corresponds with the economist's notion of labor service as a factor of production. No attempt is made to allow for different grades of labor, for changes in the intensity with which the labor is applied, for changes in the composition of the labor force, nor for the increased importance of skills and education in more recent years. We have taken labor in terms of manhours of some constant or average quality. This surely results in an understatement of the increase in manhours worked; today's worker is unquestionably better equipped intellectually, and possibly physically (due to better diets), than his temporal predecessor. It is debatable, though, whether this should be termed an increase in labor input or whether it is a form of

1. Again it was decided to use Kendrick's data, op. cit.

investment, for besides being better complemented with physical capital, today's worker is more adequately provisioned with mental capital. This quality improvement factor, then, results in an estimate of the rate of technical change that is higher than would be obtained if labor were weighted according to skill.

Capital input would ideally refer to a flow of services from the use of structures, equipment, and inventory, commensurate with the flow of labor services. The difficulty in evaluating capital services is, however, exceedingly great, in principle as well as in practice. First, a distinction must be made between employed capital and idle capital; presumably, services from the latter equal zero.[1] Second, even limiting oneself to capital actually in use, it is not obvious what method of valuation is most appropriate. It appears that the homogeneity assumption is most flagrantly violated with respect to this input, for each year's capital goods can be expected to surpass prior vintages with regard to performance, measured in any meaningful way.[2]

However, if capital input is measured by its market value, then the central problem being discussed here is defined away. When interest is centered on capital formation and technical change as conceptually distinct forces contributing to productivity increments over time, then it is important for the former to be measured net of the latter. Clearly, if all improvements in the technology are reflected in the measurement of capital input, then capital deepening will be responsible for the entire increase in output per manhour. Consequently, capital should be measured not in present value terms, but rather in a way which equates machines with the same original cost of production, deflated to allow for changes in the general price level. This is not to argue that this is "the" way to evalute the capital stock: there are several alternative methods which are equally "correct," and one's choice from among them must depend on the use to which the figures are to be put.

In the present study, then, a good case can be made for expressing capital in the following way:

$$(2.7) \quad K(t) = \int_{-\infty}^{t} I(x) e^{-\sigma (t - x)} dx ,$$

---

1. In a continuous production function of the form assumed here, the marginal productivity of a factor is always greater than zero regardless of the quantity of the factor. For this reason there is, strictly speaking, no place in the model for idle resources. To take account of the unemployment of resources which has occurred, one must admit a temporary departure from equilibrium conditions. This may be due to imperfect substitutability between capital and labor, or between different types of capital, in the short run.

2. This fact forms the basis for the analysis in Section IV, and will be discussed in greater detail there.

where σ = a deterioration factor, I = gross investment, and where K defines the concept, "physical capital." One limiting case which deserves mention is that in which σ is a discontinuous function, as given by

$$\sigma\ (t - x) = 0,\ x \leq \theta$$

$$\sigma\ (t - x) = \infty\ x > \theta$$

(2.7) then becomes:

$$(2.7a)\quad K(t) = \int_{t-\theta}^{t} I(x)dx\ ,$$

which is an expression for the gross capital stock. Measured in this way, capital formation equals gross investment less retirements (rather than depreciation); this implies a "one-hoss shay" assumption regarding the deterioration of capital goods.

In general, though, this assumption need not be made; what is essential is that σ be exclusively a reflection of wear-and-tear, and not of obsolescence. The latter is governed primarily by investors' expectations of the future rate of introduction of new techniques, and to the extent that anticipated changes in technology approximate realized changes, obsolescence will be an approximation to technical change. While for other purposes, capital should be measured in a manner which allows for obsolescence, this would be inappropriate here. To the extent that obsolescence is deducted from the measure of capital, the stock is undervalued; this in turn results in an upward bias in the estimate of the rate of innovation.

Physical capital, as defined by equation (2.7), is a summation of assets of different types, using as a common denominator the (deflated) original cost. While market value undoubtedly understates physical capital, it is equally certain that gross capital overstates it. In other words, while market valuation adds obsolescence to deterioration, gross capital valuation assumes no wear-and-tear of the asset until its retirement. What is needed is a compromise between these two measures. This compromise is provided by a capital stock series compiled by Donald G. Wooden and Robert C. Wasson.[1] Their data are based on a notion of capital which is essentially the same as the physical capital concept discussed above. This set of figures, however, relates to all capital, regardless of whether or not it is employed in production. But during a recession, when the level of employment declines sharply, idle capacity is likely to emerge as well. Consequently, there will be a

---

1. U.S. Department of Commerce, U.S. Income and Output (Washington: U.S. Government Printing Office, 1958), pp. 196-97. 1956-57 data are taken from the revised series appearing in Survey of Current Business, July 1959. See also D. G. Wooden and R. C. Wasson, "Manufacturing Investment since 1929 in Relation to Employment, Output, and Income," Survey of Current Business, XXXVI (November 1956).

discrepancy between capital items in existence and those in use, with a corresponding discrepancy between actual capacity and the flow of capital services.

Column 2 of Table 1, and Figure 2, show the time path of the ratio of capital (measured in 1954 dollars) to manhours worked in manufacturing; this series gives evidence of a strong upward trend, but is nevertheless marked by sharp cyclical instability as compared with the output per manhour figures. While the secular rise in the capital-labor ratio so measured probably represents a "real" phenomenon, in the sense that it signifies a changing composition of inputs in the aggregate production function, there is good reason to believe that the cyclical variation results, for the most part, from idle capacity, which is included in the measurement of capital. While one can, as a rule, get a reasonably good approximation to actual manhours worked, for years in which there is considerable unemployment as well as in "full" employment years, corresponding figures for the capital stock are not available.

From the standpoint of the production function, interest centers not on capital and labor in existence, but on the quantities of the factors actually employed. The latter corresponds more closely to the notion of factor proportions in economic theory. To make the appropriate adjustment in the raw capital data, some measure of idle capacity must be devised. Although some attention has been devoted to the concept of capacity and its utilization for particular industries in the post-World War II years, no attempt was made, for earlier periods, or for the aggregate, to gauge the extent to which capital was actually employed in the economy. Nevertheless, an attempt must be made to adjust the data to remove the asymmetry caused by the labor figures referring to employed labor and the capital data relating to existing capital.

One way to do this is to multiply each year's capital stock by the proportion of the labor force employed for that year. This is the procedure followed by Solow, who, while admitting that the method is far from ideal, considers it preferable to the use of the unadjusted figures.[1] While I am in agreement with Solow on this point, I think it fair to point out that this procedure is equivalent to using as inputs (1) capital in existence and (2) the labor force. This, of course, assumes that, during a recession, both factors will be unemployed to the same extent. One must remember, however, that this study is concerned not with the whole economy, but with one sector only, and there is some question as to whether one can refer to a "manufacturing labor force." On the other hand, using data for the aggregate labor force implicitly assumes that unemployment falls evenly on all sectors, clearly a heroic assumption. It seems probable that the per cent of unemployed in manufacturing is greater than the national average. Thus, if fixed proportions is a reasonable hypothesis for the short run, idle capacity will be somewhat understated by using this method, and the capital stock will be correspondingly overstated. This is offset, fortunately, by

1. Solow, op. cit., p. 314.

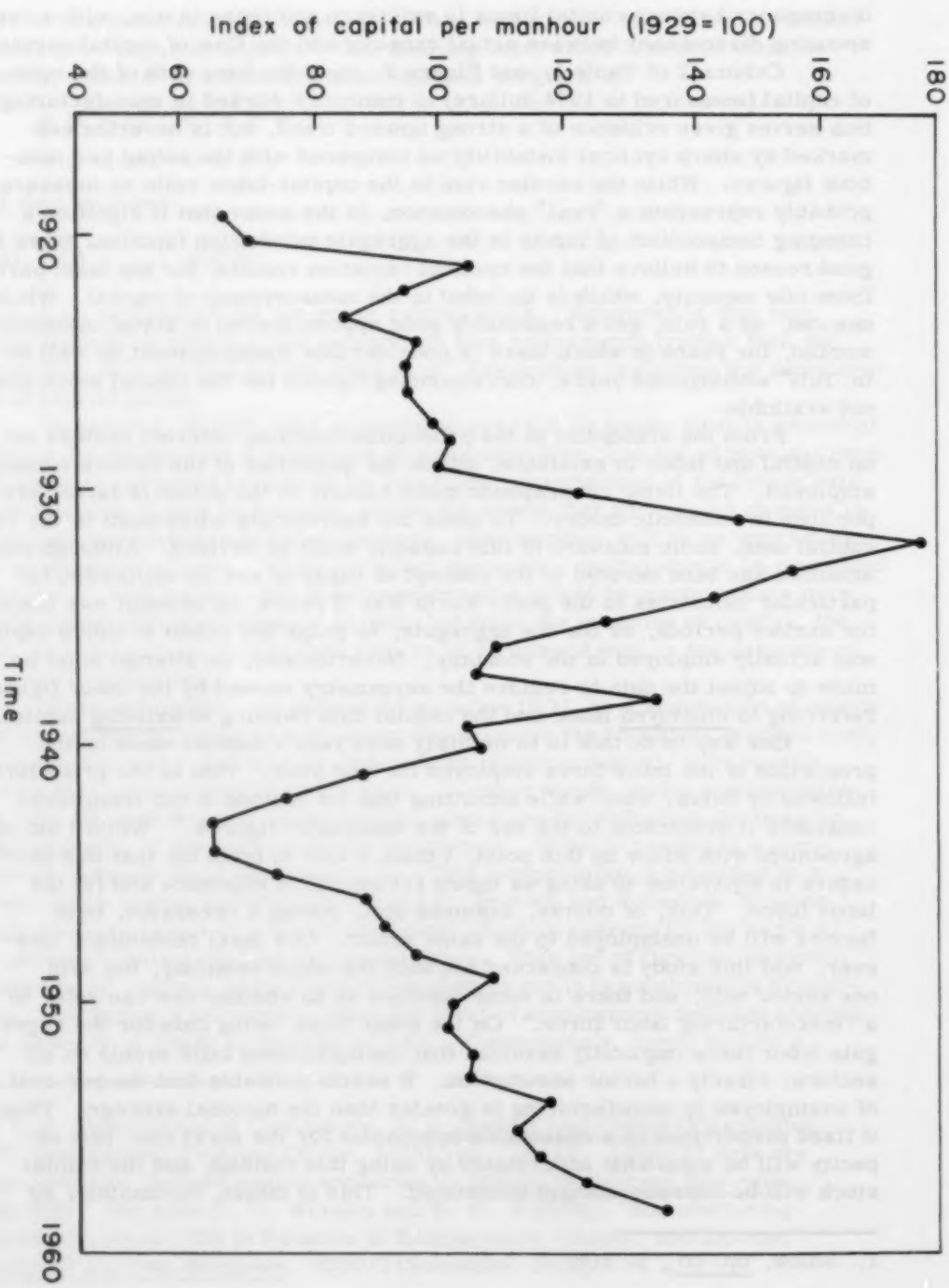

Fig. 2 — Capital per manhour in U.S. manufacturing, 1919–1958

another factor tending to err in the opposite direction.

During a recession, the labor force is likely to increase due to both secular and cyclical factors. At the same time, the level of net investment falls during such a period, possibly becoming negative for some years, resulting in little if any increase in the capital stock. Hence, a given number of employed workers constitutes a smaller percentage of labor force employment than during the preceding full employment period, whereas the capital stock has not increased proportionately. The result is an overcorrection for idle capacity, tending to balance the undercorrection discussed in the preceding paragraph. Of course, it is impossible to predict which will have the greater effect. The capital-labor data, adjusted by this procedure, are shown in Table 1, Column 3.

The fourth variable, property's share of net income originating in manufacturing (see Column 4 of Table 1), was obtained by subtracting from unity labor's share, for each year, including an imputed share for self-employed.[1] It is interesting to note that in two years, viz., 1932 and 1933, labor income exceeded total net income, indicating an apparent net loss for capital. There is some doubt, however, whether capital's share was in fact negative in these years; an alternative explanation focuses on the difference between actual depreciation of fixed capital and apparent depreciation, as represented in business records. A more "realistic" calculation of capital consumption might lead to a non-negative property income for these years. But, as stated above, the marginal productivity theory of distribution expresses only a tendency rather than an instantaneous adjustment, so that even if profits were negative, a positive marginal product of capital would be consistent with this. Because we are interested in property's share as an indication of the marginal productivity of capital, it seems reasonable to use a zero value for the two years in question.

Results of the Aggregate Study

The initial step was to calculate $\frac{\Delta A}{A}$ for each year of the interval 1919–58, as appears in Column 1 of Table 2. Next, the scale factor was determined by letting A(19) = 100.0, and the entire A(t) series was derived. This is shown in Column 2 of Table 2 and in Figure 3. Inspection of the technology path reveals that while it contains some irregularities, it is nevertheless strongly dominated by an upward trend.

Hypothesizing that A(t) can be approximated by an exponential function of time, the following equation was fit by least squares:

$$(2.8) \quad \ell n(A) = \alpha + \lambda t \ .$$

The parameters were estimated for several sub-periods, as well as for the period as a whole, and are presented in Table 3, together with the

1. Department of Commerce, U.S. Income and Output, pp. 130–31, 200; for 1956–57, see footnote 1 on p. 312.

TABLE 2

Annual Percentage Rate of Technical Change and Level of Technology in U.S. Manufacturing, 1919–1958

Year	Percentage Rate of Technical Change (decimals)	Level of Technology (index numbers, 1919=100.0)
1919		100.0
1920	.0495	105.0
1921	.0559	110.9
1922	.1429	126.7
1923	-.0241	123.6
1924	.0362	128.1
1925	.0672	136.7
1926	.0135	138.5
1927	.0196	141.2
1928	.0387	146.7
1929	.0478	153.7
1930	-.0304	149.0
1931	.0046	149.7
1932	-.0743	138.6
1933	.0886	150.9
1934	.0445	157.6
1935	.0752	169.5
1936	.0237	173.5
1937	-.0123	171.4
1938	-.0591	161.3
1939	.1153	179.9
1940	.0568	190.1
1941	.0687	203.2
1942	.0452	212.4
1943	.0491	222.8
1944	-.0144	219.6
1945	-.0507	208.5
1946	-.1163	184.3
1947	.0519	139.3
1948	.0233	198.4
1949	.0104	200.5
1950	.0923	219.0
1951	-.0145	215.8
1952	.0112	218.2
1953	.0436	227.7
1954	.0088	229.7
1955	.0726	246.4
1956	.0121	249.4
1957	.0088	251.6
1958	.0171	255.9

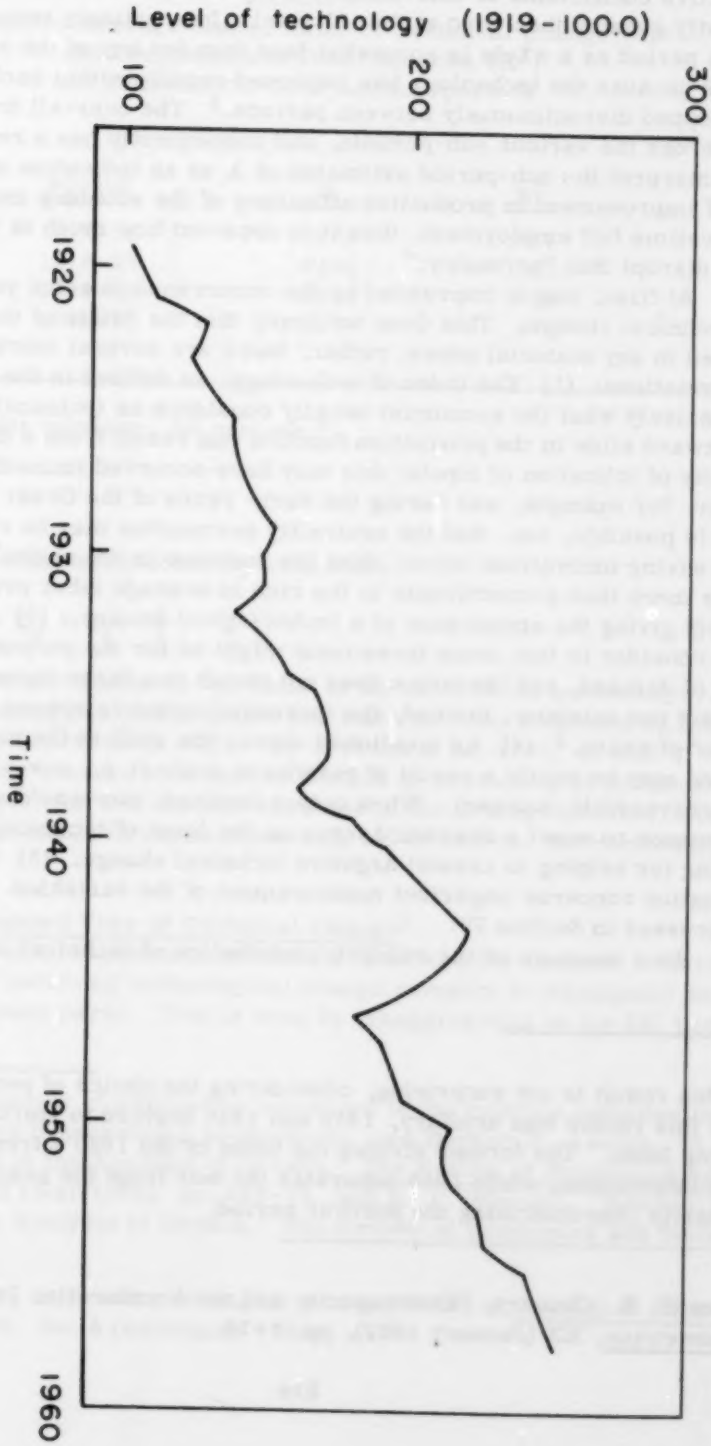

Fig. 3 — Level of technology in U.S. manufacturing, 1919–1958

respective coefficients of determination. The slope coefficients are all significantly greater than zero at the .01 level. Interestingly enough, the slope for the period as a whole is somewhat less than for any of the sub-intervals. This is because the technology has improved rapidly within each interval but has dropped discontinuously between periods.[1] The over-all trend, however, cuts across the various sub-periods, and consequently has a reduced slope. If we interpret the sub-period estimates of λ as an indication of the potential rate of improvement in productive efficiency of the economy under conditions of peacetime full employment, then it is apparent how much is lost by factors which disrupt this "normalcy."

At first, one is impressed by the occurrence in some years of negative technical change. This does not imply that the "state of the arts" has declined in any material sense; rather, there are several more plausible interpretations: (1) The index of technology, as defined in the model, is not precisely what the economist usually considers as technical knowledge. A backward slide in the production function can result from a diminished intensity of utlization of inputs; this may have occurred immediately after the war, for example, and during the early years of the Great Depression. (2) It is possible, too, that the neutrality assumption may be violated. If labor-saving innovations occur, then the increase in the capital-labor ratio will be more than proportionate to the rise in average labor productivity, possibly giving the appearance of a technological decline. (3) Another factor to consider is that some investment might be for the purpose of building ahead of demand, and therefore does not result in a large immediate rise in output per manhour; instead, the increased output is spread out over a number of years.[2] (4) As mentioned above, the shift in the production function may be partly a result of returns to scale (i.e., reversible rather than irreversible changes). When output declines, one would expect this phenomenon to exert a downward force on the level of technology, possibly creating (or helping to create) negative technical change. (5) Still another explanation concerns imperfect measurement of the variables. This will be discussed in Section IV.

As a measure of the relative contribution of technical change to

---

1. This result is not surprising, considering the choice of periods. While this choice was arbitary, 1929 and 1946 seemed to mark natural dividing lines. The former divides the boom of the 1920's from the Great Depression, while 1946 separates the war from the peaceful prosperity characterizing the postwar period.

2. See H. B. Chenery, "Overcapacity and the Acceleration Principle," Econometrica, XX (January 1952), pp. 1-28.

TABLE 3
Estimated Average Annual Rates of Technical Change (λ) and Coefficients of Determination (R^2) for Period 1919–1958 and Sub-periods

Period	λ	R^2
1919–58	.0210	.931
1919–29	.0408	.945
1929–46	.0245	.783
1946–58	.0273	.970

rising output per manhour, we can use:[1]

$$(2.9) \quad \pi = \frac{\lambda}{\xi_1},$$

where ξ_1 is estimated by least squares from

$$(2.10) \quad \ell n(q) = \xi_0 + \xi_1 t$$

with q = output per manhour. π, then, is the ratio of the percentage rate of change in technology to that in output per manhour. As indicated in Table 3, λ for the period 1919–58 is .0210. ξ_1 is found to be .2505, so that π = .84. Thus, this essay concurs with earlier studies in attributing to technical change the major share of the past increase in the average product of labor. The share attributable to capital deepening is, of course, the remainder, or .16.

3. A Disaggregated View of Technical Change[2]

This section represents an attempt to extend the analysis of Section II by refining the catch-all technological change category to distinguish among several component parts. This is done by disaggregating to the SIC two-digit

1. For a discussion of alternative measures of the relative contribution of technical change to increasing productivity, see Herbert S. Levine, "A Small Problem in the Analysis of Growth," The Review of Economics and Statistics, XLII (May 1960), pp. 225–28, and B. F. Massell, "Another Small Problem in the Analysis of Growth," The Review of Economics and Statistics (forthcoming).

2. This section, with minor changes, was published in Journal of Political Economy, LXIX, No. 6 (December 1961).

industry level and considering the 19 industry groups within the manufacturing segment of the U.S. economy, for the period 1946–57. Disaggregation yields additional information regarding productivity change, e.g., revealing the dispersion around the average of industry rates of technical change, thus indicating the extent to which shifts in the production function have been evenly spread throughout the manufacturing sector or concentrated in certain industries. It also permits us to form an estimate of the importance of the aggregation problem.

A Disaggregated Model

Disaggregation enables a distinction to be drawn between the reallocation of resources among industries and the improvement of efficiency within industries; both factors contribute to the shift in the aggregate production function, although, strictly speaking, only the latter should be termed innovation in the usual sense. The model outlined above can be extended to express intraindustry improvements in the state of the arts as a weighted sum of the improvement factor for the individual industries. This component can be subtracted from aggregate technical advance, yielding an interindustry component residually.

First, write the percentage rate of increase in output for industry i as follows:

$$(3.1) \quad \frac{\dot{Q}_i}{Q_i} = \frac{\dot{A}_i}{A_i} + \beta_i \frac{\dot{K}_i}{K_i} + (1 - \beta_i) \frac{\dot{L}_i}{L_i} \quad .$$

Next, define:

$$w_i^L = \frac{L_i}{L} \quad , \quad w_i^K = \frac{K_i}{K} \quad .$$

It then follows that

$$(3.2) \quad \frac{\dot{L}_i}{L_i} = \frac{\dot{L}}{L} + \frac{\dot{w}_i^L}{w_i^L}$$

and

$$(3.3) \quad \frac{\dot{K}_i}{K_i} = \frac{\dot{K}}{K} + \frac{\dot{w}_i^K}{w_i^K} \quad .$$

Now, the aggregate percentage rate of increase in output is easily seen to be:

$$\frac{\dot{Q}}{Q} = \frac{\Sigma \dot{Q}_i}{Q} \quad ,$$

which can be written:

$$(3.4)\quad \frac{\dot{Q}}{Q} = \Sigma \left[\frac{\dot{Q}_i}{Q_i} \cdot \frac{Q_i}{Q} \right] .$$

Substituting (3.1), (3.2) and (3.3) into (3.4), we can write:

$$(3.5)\quad \frac{\dot{Q}}{Q} = \Sigma \frac{Q_i}{Q} \left\{ \frac{\dot{A}_i}{A_i} + \beta_i \left[\frac{\dot{K}}{K} + \frac{\dot{w}_i^K}{w_i^K} \right] + (1 - \beta_i) \left[\frac{\dot{L}}{L} + \frac{\dot{w}_i^L}{w_1^L} \right] \right\} .$$

Noting that:

$$\Sigma \beta_i \frac{Q_i}{Q} = \beta ,$$

and that:

$$\Sigma (1 - \beta_i) \frac{Q_i}{Q} = 1 - \beta ,$$

we can write, rearranging terms:

$$(3.6)\quad \frac{\dot{Q}}{Q} - \beta \frac{\dot{K}}{K} - (1 - \beta) \frac{\dot{L}}{L} = \Sigma \frac{Q_i}{Q} \left[\frac{\dot{A}_i}{A_i} + \beta_i \frac{\dot{w}_i^K}{w_i^K} + (1 - \beta_i) \frac{\dot{w}_i^L}{w_i^L} \right] .$$

Noting that the expression on the left is aggregate technical change as given by (2.2e), it follows that:

$$(3.7)\quad \frac{\dot{A}}{A} = \gamma_1 + \gamma_2 + \gamma_3 ,$$

where

$$(3.8)\quad \gamma_1 = \Sigma \frac{Q_i}{Q} \cdot \frac{\dot{A}_i}{A_i} ,$$

$$(3.9)\quad \gamma_2 = \beta \Sigma \frac{f_i^K}{f^K} \dot{w}_i^L ,$$

$$(3.10)\quad \gamma_3 = (1 - \beta) \Sigma \frac{f_i^L}{f^L} \dot{w}_i^L ,$$

and where f_i^K and f_i^L denote, respectively, the marginal product of capital and of labor in industry i, and where the absence of a subscript denotes the aggregate.

Let us look at these three terms. γ_1 can be regarded as intraindustry technical change. It consists of a weighted average of the rates of technological advance in the individual industry groups. In other words, this term results from improvements in efficiency within the industry groups, and excludes improvements resulting from a reallocation of resources among industries. It thus represents innovation in the strict sense of the word.

Subtracting γ_1 from aggregate technical change leaves us with interindustry technical change, $\gamma_2 + \gamma_3$, arising from an interindustry shift of resources. Either labor or capital may shift from industries where its marginal productivity is low to those where it is high, resulting in an improvement in aggregate technology over and above that attributable to improvements within the industry groups. This component of technical change is introduced by the process of aggregation, as can be seen from the following example. Suppose we view technical change as occurring in two stages. In the first stage, production functions are changed in some or all of the industries, with a resulting increase in output, but no inputs are permitted to shift among industries. This will typically result in a disequilibrium situation. In the second stage, inputs are shifted among industries so as to restore equilibrium. This reallocation of inputs results in a further increase in output, over and above the direct increase achieved in the first stage. But, by assumption, there is no change in the techniques of production of any industry in the second stage; thus $\gamma_1 = 0$, for it consists of a sum of terms all of whose value is zero. If the aggregate quantities of factors are held constant, then, from equation (2.3), the second stage increase in output must appear as an improvement in aggregate technology.

While the interindustry component is technical change as defined in equation (2.3), it should be possible to distinguish this from more basic changes in the technological structure of the economy, often referred to as "innovational" in nature. It is the latter which we would like to measure, but it is a mixture of the two which appears in our estimates. We can view the relative importance of the former component, i.e., interindustry technical change, as a measure of the significance of the "aggregation problem" since it indicates what fraction of aggregate technical advance results from the process of aggregating over industries. In other words, the macro variables in the model are not exact counterparts of the corresponding micro variables.

It should be noted that even the intraindustry component of technical change is not really "pure." The present section considers disaggregation to the SIC two-digit industry level. Further disaggregation to, say the four-digit level, would result in a further separation of γ_1 into three components, as in equation (3.7). This process could, of course, be continued to the point at which we have observations on the technological structure of the individual firm, or even plants within a firm. We would undoubtedly find that at each level of aggregation, terms corresponding to our interindustry technical change would enter the relationship.

While such a study might be desirable, it should be pointed out that the limitations of the model presented in equations (2.1)–(2.4) make such disaggregation inadvisable. In particular, assuming the neutrality of shifts in the production function appears somewhat tenuous at any level of disaggregation beyond that considered in the present paper; some may question it even here.

Interindustry Technical Change

The movement of factors among industries can be identified as a manifestation of a dynamic adjustment process. Without specifying additional behavioral equations, it is not possible to describe the nature of this adjustment in any great detail. It almost certainly comprises a complex set of forces operating simultaneously to produce the resultant observable effect. At least three distinct types of forces can be enumerated: (1) The reallocation of resources during the period may, to some extent, represent a long, continuous process of readjustment to an initial disequilibrium. (2) Alternatively (or additionally), the shifting of resources may represent a process of adjustment to a continuously shifting equilibrium position. (3) Finally, to the extent to which it is labor that does the shifting, it may merely be a manifestation of interindustry differences in the "quality" of the labor force. Each of these will be considered briefly.

According to the first explanation, a certain degree of disequilibrium existed at the beginning of the period. The return to either capital or labor (or both) was higher in some industries than in others, and the resource moved towards those uses with the higher returns. The process of adjustment, however, may have occurred gradually throughout the period rather than instantaneously,[1] the same industries continuing to receive resources while other industries continued to lose resources. In this case, the ranking of industries on the basis of marginal productivity would not have varied significantly during the period, although the extent of the discrepancy would have been reduced. According to this explanation, $\gamma_2 + \gamma_3 > 0$ implies a real improvement in the efficiency of resource allocation during the period.

The second explanation of interindustry technical change hypothesizes a dynamic process of adjustment to an equilibrium which itself was shifting during the period. Thus, in this case, there need be no secular improvement in allocative efficiency in the sense implied above. Either of three forces can disturb the equilibrium. First, changes in the pattern of demand may change the relative values of the goods produced by the various industries, resulting in interindustry differentials in the wage rate or return to capital. Alternatively, differential rates of technical change may increase the returns to factors in some industries relative to others, again creating an impetus to factor mobility. Finally, an interindustry discrepancy in marginal productivities of factors could also result from random shocks affecting

1. This might result from institutional barriers to short-run factor mobility.

each industry's output, and the adjustment occasioned by this phenomenon would appear as interindustry technical change even though it did not represent a real gain to the economy. Such random variability in industry output will also result in an overstatement of aggregate technical change, as calculated in Section II, and can thus be regarded as another component of the aggregation problem. For example, if each industry's production function is subject to a random disturbance, we would have

$$(3.11) \quad Q_i = A_i(t)\, f_i(K_i, L_i)\, u_i(t) ,$$

where u is stochastic. (Note that u must enter the relationship multiplicatively in order for the operations in equation [3.2] to be performed.) Now, if $E(u) = 1$, and if $E[u_i(t)\, u_j(t)] = 0$ and $E[u_i(t)\, u_i(t-1)] = 0$, then intraindustry technical change would be unaffected by the substitution of a stochastic production function for an exact one.

The third explanation of interindustry technical change centers on the heterogeneity of labor. Because equations (2.1)–(2.4) implicitly assume a homogeneous labor force, it is apparent that any real interindustry difference in the average skills and abilities of the workers will appear as a difference in the technology levels of the industries in question. The marginal productivity of more highly skilled labor will exceed that of less skilled workers, and there will be a corresponding difference in the wage rates. If some industries tend to employ higher quality labor, then the average wage paid by these industries will also be higher. Now, to the extent that labor moves into higher paying industries, this movement will show up as interindustry technical change. This is, of course, a form of technical change, broadly defined; but, more precisely, it results from an increase in the quality of the labor force. That is, by learning new skills, a worker is enabled to move to a better job (perhaps in a new industry) where he will receive greater pay to compensate for his increased marginal contribution to output. This improvement in labor force quality, to the extent that it is manifested in interindustry mobility, is going to be included in the Section II measure of technical change.[1]

---

1. One may question the consistency of interindustry disequilibria with the above assumption that factors are competitively priced in the large. In this regard, the present section can be viewed as an application, at the industry level, of assumptions previously made for the aggregate. It is certainly legitimate to ask whether the assumed competitive pricing of factors in each industry implies, as a result of the empirical work performed below, that aggregate factor markets are not perfect. This, in itself, should be of some interest, if only as an empirical test of one of the assumptions underlying the earlier analysis.

However, I feel that the presence of interindustry technical change is

Empirical Results

The initial step was to calculate, for each industry, the percentage rate of technical change, $\dot{A}_i/A_i$, for each year of the interval 1946–57. Next, the technology level for the initial year in each industry was arbitrarily set equal to unity, and the technology levels for the following years were expressed as index numbers, employing equation (2.4). For this procedure, annual observations were needed on output, manhours worked, capital input, and property's share of the product, as in Section II.

As a measure of each industry's output, we used the Federal Reserve Index of Industrial Production.[1] This procedure raises the question as to how industry outputs can be summed to obtain aggregate output, needed for the calculation of γ_1. To surmount this difficulty, each industry's output series was weighted by the value added in the industry in 1954.[2] Thus, if the ratio of external purchases to total output remains constant for each industry, as assumed above, this procedure amounts to using as weights the relative prices of the output produced by the industries. As in any construction of an aggregate output series, it is assumed that these relative prices remain constant.[3, 4]

But, one may ask, if relative prices remain constant, then how is equilibrium restored? If the marginal productivities of labor and capital were higher in one industry than in others, and if the production function

---

consistent with the earlier assumption if this assumption is re-interpreted. For example, there may be institutional barriers to interindustry resource mobility in the short run, whereas, in the longer run, the barriers can be surmounted, at least in part. Certainly the short-run heterogeneity of labor is one manifestation of this. Then, in each industry, factors may continuously be competitively priced, but the allocation of factors among industries will be nonoptimal. The aggregate β must then be interpreted as a weighted average of the industry β's. One cannot, in this context, speak of the marginal productivity of a factor in the aggregate, except as a form of average.

1. U.S. Department of Commerce, Business Statistics (U.S. Government Printing Office: Washington, 1959), pp. 8–10

2. National Industrial Conference Board, Economic Almanac (New York. Crowell, 1958), p. 190.

3. Thus, an increase in an industry's index of production will be reflected in the aggregate as an equal percentage increase in the industry's value added.

4. Aggregate technical change, calculated on the basis of this measure of aggregate output, does not differ materially from the measures presented in Section II.

were linear homogeneous (as assumed), then an interindustry shift of factors would not correct the discrepancy in factor returns, and factors would continue to shift until they were all in one industry. However, one would expect that as the output of the progressive industry rises, its price would decline such that the marginal value productivities would begin to decline, restoring equilibrium.

If price changes are admitted, how is our analysis affected? For one thing, prices are likely to fall in industries which experience a rapid rate of technical change. Hence, industries which have had the most rapid technical change during the period will have lower end-of-period prices (relative to initial prices) than less progressive industries. By using mid-period (1954) weighting, our procedure thus tends to understate interindustry technical change for early years and overstate it for later years. The reverse holds for the components of interindustry technical change.

Turning to the other variables, labor input was measured in manhours, as in Section II, using Department of Commerce data.[1] Persons engaged includes the labor time of self-employed persons as well as employees. Property's share was obtained by subtracting from unity the ratio of labor income[2] (including an imputed share for the self-employed)[3] to net income originating in the industry.[4]

Regarding capital input, it was assumed that the proportion of "physical" capital in each industry was equal to the industry's share of book value capital. The latter was obtained from Treasury data on corporations,[5] was adjusted to allow for non-corporate capital,[6] and was subsequently multiplied by the ratio of physical capital to book value capital in manufacturing as a whole. The Section II data, adjusted for idle capacity, were used for the physical capital series.

Hypothesizing that $A_i(t)$, the level of technology in industry i at time t, can be approximated by an exponential function of time, the following

---

1. U.S. Department of Commerce, U.S. Income and Output, op. cit., p. 214; Business Statistics, op. cit., pp. 67-69 (1956-57 data are taken from the revised series appearing in Survey of Current Business, July 1959).

2. Commerce, U.S. Income and Output, op. cit., p. 200. (For 1956-57, see footnote 1 above.)

3. Employee compensation was multiplied by the ratio of persons engaged to employees.

4. Commerce, U.S. Income and Output, pp. 130-31. (For 1956-57, see footnote 1 above.)

5. Difference of gross capital assets (including inventories) and accumulated depreciation reserves. U.S. Treasury Department, Statistics of Income, 1947-57 (Washington: U.S. Government Printing Office).

6. By multiplying by the ratio of total sales to corporate sales.

equations were fitted by least squares:

$$(3.12)\quad \ell n\,[A_i(t)] = \alpha_i + \lambda_i t \qquad i=1, \ldots, 19$$

The estimated values of λ are presented in decreasing order in Table 4, together with the respective coefficients of determination (R^2), indicating how close a fit is provided by an exponential approximation.

As would be expected, for the most rapidly improving industries, a large proportion of the variation in $\ell n(A)$ is explained by time, whereas R^2 is below .600 in five of the seven least progressive industries. It is comforting that λ is positive for all industries, and not suspiciously large for any. Aside from this, the degree of dispersion is evidence that the percentage rate of improvement in technology has certainly not occurred evenly throughout the sector.

As a check on this procedure, the ratio, ω, of end-of-period technology to beginning-of-period technology was computed for each industry; the ordering based on ω was compared with that based on λ, and Spearman's rank correlation coefficient was computed. The value of Spearman's coefficient was found to be .68, with a standard error .24, for a sample of size 19. If Spearman's coefficient is assumed to be normally distributed for a sample of this size, then there is a significant positive correlation between the two orderings at the .01 significance level, employing a one-tailed test.

At first glance, one is perhaps surprised to find lumbering at the top of the list in Table 4, and primary and fabricated metal products at the bottom. (Ranking on the basis of ω confirms these results.) The peculiarity of these results could be an indication of non-neutral technical change in these industries. For example, were technical change predominantly capital-saving during the period, the rate of technical change would be overstated by the procedure employed here. On the other hand, labor-saving improvements in technology would tend to give a downward bias to the measure of technical change; this may have been the case in the metal industries. Another explanation for the unexpected results may be a bias in the measurement of capital. Book-value capital, duly adjusted, is used as a measure of capital input. But, if an industry experiences a decline during the period in the ratio of employed capital to book-value capital (as a result, e.g., of increasing idle capacity), the rate of technical change will be understated accordingly.

The next step was to calculate interindustry and intraindustry technical change. Table 5 summarizes the results of this procedure.[1] Intra-industry technical change, shown in Column 1, reaches a high of .036 in 1954 and a low of .002 in 1952. These correspond to the high and low years,

1. Because of shortcomings in the data, it was necessary to drop the years 1946–48. Moreover, the resulting figures were converted to 3-year moving averages, with a consequent loss of two additional years, 1949 and 1957.

TABLE 4

Estimated Average Percentage Rate of Technical Change (λ), Coefficient of Determination (R^2), Initial Technology Level (A) and Average Value of Property's Share ($\bar{\beta}$) in Manufacturing Industry Groups

	Industry Groups	λ	R^2	A (1949)	$\bar{\beta}$
1.	Lumber and wood products	.0377	.860	.754	.169
2.	Electrical machinery	.0369	.821	1.286	.221
3.	Chemicals and products	.0346	.926	1.447	.393
4.	Stone, clay, and glass products	.0248	.890	1.291	.275
5.	Printing and publishing	.0244	.886	1.817	.158
6.	Transportation equipment	.0239	.857	1.525	.273
7.	Pulp, paper and products	.0234	.773	1.450	.303
8.	Machinery, except electrical	.0197	.927	1.538	.212
9.	Petroleum and coal products	.0189	.702	.432	.557
10.	Textile mill products	.0159	.837	.954	.164
11.	Food and kindred products	.0141	.932	1.667	.230
12.	Leather and leather goods	.0106	.604	1.185	.116
13.	Rubber products	.0098	.369	1.557	.221
14.	Furniture and fixtures	.0097	.767	1.176	.142
15.	Instruments and related products	.0097	.546	1.528	.221
16.	Apparel and related products	.0086	.861	1.259	.091
17.	Tobacco manufactures	.0078	.408	.794	.532
18.	Primary metal products	.0043	.046	1.451	.268
19.	Fabricated metal products	.0025	.025	1.498	.170

TABLE 5

Technical Change, Intraindustry and Interindustry

Year	(1) Intraindustry Technical Change γ_1	(2) Total Technical Change $\gamma_1 + \gamma_2 + \gamma_3$	(3) Ratio of Intraindustry to Total $\frac{\gamma_1}{\gamma_1 + \gamma_2 + \gamma_3}$	(4) Interindustry Technical Change $\gamma_2 + \gamma_3$
1950	.023	.027	.85	.0044
1951	.021	.031	.68	.0105
1952	.002	.015	.13	.0130
1953	.007	.022	.32	.0145
1954	.036	.040	.90	.0040
1955	.020	.029	.69	.0091
1956	.024	.029	.83	.0053
Mean	.019	.028	.68	.0087

respectively, for the aggregate, shown in Column 2. The latter, however, is more stable, so that the ratio of the two, shown in Column 3, is subject to large fluctuations, between .13 in 1952 and .90 in 1954. Interindustry technical change is rather more stable, fluctuating between .0040 and .0145.

While some interest attaches to the annual fluctuations in these variables, their mean values for the period as a whole are, perhaps, more interesting. The arithmetic mean of the annual rates of technical change for the period was .028, while the average intraindustry change amounted to .019. The difference, .009, accounted for 32 per cent of the total, a surprisingly high figure. Thus nearly one-third of the technical change resulted from an interindustry shift of resources. As Section II attributes 84 per cent of the increase in output per manhour to changes in technology, these new results suggest that approximately one-fourth of the increase in average labor productivity may have resulted from a shift of resources among industries.

As noted above, interindustry technical change can result from the movement either of capital or of labor from industries where its marginal product is low to those where it is higher, corresponding to γ_2 and γ_3, respectively. These two components appear in Table 6. γ_3, shown in Column 1, varies between .004 in 1954 and .0027 in 1951. γ_2, which appears in Column 4, fluctuates between .0032 in 1950 and .0140 in 1953.[1] The ratio of γ_3 to total interindustry change, Column 3, varies between .03 in 1953 and .27 in 1950, with a mean value of .14. The mean value of γ_3 is .0012 as compared with .0075 for γ_2. Thus, shifts of labor accounted for only about one-seventh of the total interindustry shift of resources, the remainder being attributable to capital mobility. This fact has particular significance when we recall that γ_3 includes labor force quality improvement.[2]

Interindustry Gap

The data indicate a high degree of resource mobility during the period 1950–56, particularly with regard to capital. One might ask whether this has influenced the rates of technological advance of the various industries. In particular, it might be interesting to inquire whether there has been a narrowing of the dispersion among industries with regard to the level of technology. Or, on the contrary, have the best grown better while the less technologically developed industries lagged behind?

---

1. One might expect γ_2 to have been positively correlated with the dispersion of the rates of return to capital, but it turns out that this was not so. A regression of γ_2 against $\sigma^2_{f_i}K$, the variance of the f_i^K, was not significant. However, this may be due in part to the relative constancy of $\sigma^2_{f_i}K$ during the period.

2. In other words, if part of the interindustry technical change resulted from a shift of labor to more highly skilled jobs, then labor mobility as a response to disequilibria must have been less than 14 per cent of the total.

TABLE 6

Breakdown of Interindustry Technical Change

Year	(1) Interindustry Change Due to Labor Shifts γ_3	(2) Total Interindustry Change $\gamma_2 + \gamma_3$	(3) Proportion of Total Interindustry Change Due to Shifts in Labor $\frac{\gamma_3}{\gamma_2 + \gamma_3}$	(4) Interindustry Change Due to Shifts in Capital γ_2
1950	.0012	.0044	.27	.0032
1951	.0027	.0105	.26	.0078
1952	.0016	.0130	.12	.0114
1953	.0005	.0145	.03	.0140
1954	.0004	.0040	.10	.0034
1955	.0006	.0091	.07	.0085
1956	.0011	.0053	.21	.0042
Mean	.0012	.0087	.14	.0075

To answer these questions, however, the notion of an absolute level of technology must be defined. Measures of the technology level of industries have thus far been expressed only as index numbers, with the initial year in each case set equal to unity. To permit interindustry comparisons of the technology level, it is necessary to make an additional assumption, viz., that the production possibilities open to each industry can be characterized by a Cobb-Douglas function (though the parameters need not be the same for any two industries). Specifying the form of the function, together with the set of assumptions made above, enables rewriting (2.1):

$$(3.13)\quad Q_i(t) = A_i(t)\, K_i(t)^{\beta_i} L_i(t)^{1-\beta_i} ,$$

where the subscript designates the industry. If (3.13) is solved for A, the result is

$$(3.14)\quad A_i(t) = \frac{Q_i(t)}{K_i(t)^{\beta_i} L_i(t)^{1-\beta_i}} .$$

Assuming the production function to be a Cobb-Douglas implies that the ratio of capital's marginal product to its average product remains constant regardless of the point on the curve at which the industry is operating; moreover, for the neutrality assumption to be satisfied, the curve must be shifting isoelastically, so that the capital coefficient is invariant with respect to time. Consequently, it is necessary to reinterpret the earlier assumption of competitive pricing in factor markets to imply not that each year's value of property's share is an estimate of the coefficient for that year, but that it is an estimate of the unique coefficient (i.e., unique for each industry) which we shall take to be the arithmetic mean of the observations, that is,

$$(3.15)\quad \beta_i = \frac{1}{T} \sum_{t=1}^{T} \beta_i(t) .$$

The β's are shown in Column 4 of Table 4.

Having defined the notion of absolute technology, one can compare industries on this basis (see Table 4, Column 3). The level of technology of an industry is partly a reflection of both the state of the arts existing in the industry and the extent to which firms are using the most efficient techniques. Partly too, though, it is affected by the elasticity of demand for the industry's products. For if an industry has had rapid technological progress in the past, then the extent of the price decline occasioned by this will influence the value of A_i. Of two industries with the same technological possibilities, the one with the greater monopoly power is likely to have a higher value of A_i, due to its ability to squeeze greater-valued output from its inputs. Even if all industries were perfectly competitive, though, there would

be no reason to expect $A_i = A_j$, for all i and j, unless we also have $\beta_i = \beta_j$.[1]

If interindustry dispersion in the technology level is measured as the variance of A, then a comparison of the variance at the beginning of the period with that at the end of the period will indicate whether the gap among industries has grown wider or more narrow. An F-test can be used to test for the significance in the difference between the two variances, letting the F statistic equal the ratio of end-of-period variance to beginning-of-period variance. A value of F significantly less than unity indicates that the gap has narrowed, while if F significantly exceeds one, it is evidence of a widening of the gap. A two-tailed test is appropriate.

The variances were computed and the F ratio was found to be $F = 1.33$ where there are 18 degrees of freedom both for numerator and denominator. This value of F is not significant at the .05 level, thus there is no evidence of a change in the dispersion of industries with respect to their technology coefficients.

A somewhat different test consists of regressing λ, the estimated average annual rate of technical change, against $A(t_1)$, the initial technology level, again using a two-tailed test. This shows whether there is any connection between the industries which experienced the most rapid technological advance during the period and those with the highest initial technology coefficient. The results of this regression are (with the t-ratio in parentheses):

$$\lambda = .036 - .014\, A(t_1)$$
$$(-.52)$$

The slope coefficient is not significant at the .05 level, suggesting the absence of a relationship between an industry's technology level and its rate of technological advance.

4. Interaction Between Technical Advance and Investment[2]

The question arises as to whether anything like the observed expansion in output per manhour would have occurred in the absence of capital

---

1. A_i is not a simple measure of valued output per unit value input, which would be $\dfrac{Q_i}{f_i^K K_i + f_i^L L_i}$.

2. This section, with minor changes, appeared in Metroeconomica, XIV, Nos. 1 and 2 (1962).

formation.[1] Although much technical improvement consists of organizational changes in the means of combining existing factors and does not require additional inputs, a substantially greater proportion of technological developments is probably embodied in new capital goods. Hence there is strong reason to believe that the rate of innovation is influenced by the rate of capital formation. Probably a more direct relationship can be found between changes in the technology and gross investment, for innovations are frequently incorporated in capital items which represent replacements of existing plant and equipment. Therefore, even in the absence of net investment, there can be a rapidly improving technology as long as obsolete items are being replaced by more modern ones.[2]

This is not to deny the significance of the distinction between a movement along a fixed production function and a shift from one curve to another. One must grant that such a separation of causal forces is meaningful unless he is prepared to abandon the production function concept altogether. But visualizing the growth process exclusively in these terms has the disadvantage of obscuring the interrelationship between the two sources of productivity increase. An additional question that might be asked is, how much of the observed rise in output per head would have occurred in the absence of capital formation? Suppose a distinction is made between "autonomous" and "induced" shifts in the production function. The former can occur without the aid of investment, while the latter require new plant or equipment expenditure as a vehicle. Determining the interaction between investment and technical change amounts to estimating the relative significance of these two varieties of technical change.

A Generalized Model

Induced innovation can be viewed as occurring in two stages. First, a new technique is invented and adopted by a limited number of firms (possibly only one), and second, the new process becomes more widely known and diffuses throughout the industry. But the flow of inventions to a society is a continuous process; new techniques are constantly being developed, so that the methods which are most effective today may become obsolete by tomorrow. What one sees at any point in time is an array of processes

---

1. This issue is discussed in my "Investment, Innovation, and Growth," Econometrica, 30 (April 1962), pp. 239–52. See also Leif Johansen, "Substitution Versus Fixed Production Coefficients in the Theory of Economic Growth: A Synthesis," Econometrica, 27 (April 1959), pp. 147–76; and Robert Solow, "Investment and Technical Progress," in Mathematical Methods in the Social Sciences, K. J. Arrow, S. Karlin, and P. Suppes, editors (Stanford: Stanford University Press, 1959).

2. While it may be objected that equipment can often be improved through modification, without necessitating replacement, it should be clear that such modification is almost certain to involve investment.

existing simultaneously, representing a continuum of technologies. Not only will the level of technology vary among firms, but there is likely to be some heterogeneity within a single plant as well.

It seems reasonable to suppose that nearly all additions to the capital stock will incorporate the latest (or nearly the latest) know-how, so that a firm which has just expanded its facilities or renovated its plant can be described as operating along a "best-practice" production function, at least with regard to the new installation.[1] The technology associated with a complex of capital will be a function of the year in which the items were built, and this will be a non-decreasing function of time (assuming technological knowledge does not deteriorate). In other words, each vintage of capital is a different input, and while there is a common denominator by which the entire stock can be aggregated, it should be recognized that more recent vintages represent a higher level of know-how, and hence are more productive than their forebears.

As the state of the arts continues to improve, the relative position of a firm with a given vintage of capital (and the technique associated with it) will deteriorate, i.e., in George Terborgh's words, the equipment will develop "operating inferiority" relative to the most recent models.[2] Were all firms to adopt new processes instantaneously, then the best-practice technology would prevail throughout the industry and would be the same for each firm. In fact, however, firms do not continuously maintain the most modern level of technology. Because it is uneconomic to scrap existing equipment each time a modification is introduced, there is a lag between the invention or discovery of a technique and its widespread usage.[3]

---

1. The distinction between best-practice and average production functions appears in a paper by Anne Gross, "The Technological Structure of the Cotton Textile Industry," in Leontief, Studies in the Structure of the American Economy (New York: Oxford University Press, 1953). See also K. Maywald, "The Best and the Average in Productivity Studies and in Long-Term Forecasting," Productivity Measurement Review, 9 (May 1957).

2. See George Terborgh, Dynamic Equipment Policy (New York: McGraw Hill, 1949).

3. The length of this lag is determined by several considerations. From the point of view of the firm, the relevant consideration is the cost of producing a given output with its existing equipment, as compared with the cost of producing the same output with best-practice equipment (including the purchase price). The cost differential between new and old equipment is a function of both deterioration and obsolescence; of course, obsolescence, in turn, is determined by the rate of change in the best-practice technology. The best-practice technology thus influences the actual technology, not only by determining the efficiency of new capital goods at each point in time, but also by influencing the replacement policy pursued by firms. One would

Now the actual level of technology of an industry will consist of a weighted average of the technologies associated with the individual firms. If each firm's technology represents what was (or is) a best-practice technique, then the industry level will be an average of the best-practice techniques associated with all the vintages of capital which are represented; only one of these (this year's) is still best, however. The weight of each technique in the average depends on the amount of capital which incorporates it. In other words, the proportion of the total stock which consists of vintage $\underline{x}$ determines the weight attached to technique $\underline{x}$ in computing the average efficiency of the industry. The quantity of vintage $\underline{x}$ is a function of investment in the interval during which this vintage was best (subtracting from this the amount of vintage $\underline{x}$ which has been withdrawn from service).

This provides a framework for a theory of technical change. The level of technology prevailing in an industry, i.e., the average level, will be determined by the various individual technologies and the weights attached to each; that is, it will depend on the state of the arts in each of the preceding years and how much investment took place in each year (deducting capital which has been scrapped), relative to the total stock. At an accelerated rate of invention, each vintage will be better than it would have been at a slower rate, so that the average will be correspondingly better. But, additionally, if the age composition of the capital stock is shifted such as to favor newer types, the average will be improved even though none of the individual technologies has been changed. This leads to the conclusion that, following a period of greater than average investment, the capital stock will be heavily weighted with new vintages, and for this reason the average quality will be raised.

If A(t) is expressed as an exponential function of time, then the

---

think that a more rapid rate of improvement in the state of the arts would provide an incentive for capital goods producers to shorten the intended durability of equipment. If it is uneconomic to operate a machine for more than 10 years, then why build one to last for 20? Short-lived assets could be replaced by more efficient ones whenever new developments in the technology occurred, at a smaller cost to society, assuming that a less durable good costs less to build. Indeed, this factor could, in turn, result in a decline in the optimal service life. Note, however, that there are likely to be "economies of durability" in the production of capital goods; i.e., while the cost of an asset may well be an increasing function of the durability built into that asset, one would expect the second derivative to be negative, so that, e.g., doubling the intended service life would less than double the cost. On this point, see R. C. Blitz, "Capital Longevity and Economic Development," American Economic Review, XLVIII (June 1958), pp. 313–22. For a somewhat different consideration, cf. A. S. Manne, "On the Timing of Development Expenditures and the Retirement of Military Equipment," TEMPO SP-88.

preceding discussion would seem to suggest that there is a positive relationship between residuals from the technology trend and the ratio of gross investment to capital input in the immediate past.[1] The extent of this relationship can be taken as an indication of the importance of induced innovation, while the remaining part of technical change can be considered autonomous.

Another way to view this relationship is in terms of two new concepts: the average marginal productivity of capital (AMP) and the marginal marginal productivity of capital (MMP). The AMP is the relevant concept when homogeneity of capital is assumed. It is the marginal contribution to output of a unit of capital with the same efficiency as existing assets. If capital is heterogeneous, then AMP applies to capital whose efficiency is an average of the efficiencies of existing assets. Contrasted with this, MMP is the marginal contribution of a unit of <u>new</u> capital, taking account of the superior technology associated with with this new asset. If capital is homogeneous, MMP and AMP are identical; otherwise, MMP will exceed AMP. In other words, a dollar's worth of new machinery will produce more than a dollar's worth of average machinery.

The model presented in Section II was based on the AMP notion of marginal productivity, and the indices of technology were derived on this basis. Now, however, the question is being asked as to whether AMP does not understate the contribution of investment to increases in productivity. If innovations are embodied in new capital goods, the increase in output resulting from a unit of investment will instead be given by MMP. If this is the case, the deviations from trend in the level of technology (as estimated by the Section II model) will be positively correlated with the investment-capital ratio.

Further, it might be hypothesized that the strength of the association between these two variables diminishes over time, i.e., the contribution made by investment of x years ago is greater than that of $x + 1$ years past. The most recent investment, then, can be expected to have the greatest effect, but all investment is taken into account. The level of technology will deviate positively from its trend following a period of high investment (relative to the capital stock), while residuals will be negative following a group of low-investment years.

Now, if we assume that the level of technology can be expressed as an exponential function of time, as in Section II, we can write,

$$(4.1) \quad \ell n\, A(t) = \alpha + \lambda t + u(t) \ ,$$

where u = the residual. The above discussion suggests that u(t) can be expressed as a distributed lag,

$$(4.2) \quad u(t) = \zeta_0 + \sum_{i=1}^{\infty} \zeta_i \, \ell n\, I(t - i)$$

1. Using deviations from the trend in A(t) amounts to the same thing as using A(t) as the dependent variable and time as an additional independent variable.

where I = the gross investment–capital stock ratio. The dependent variable is influenced by all investment, but the coefficients decrease as i increases, expressing the fact that the effect of a unit of investment varies inversely with the length of the lag. A special form of lag is the Koyck distributed lag, which specifies that after a certain point, m, the relationship can be described by a decaying geometric series.[1] If equation (4.2) is assumed to be of this form, with m = 0, it can be rewritten

$$(4.3) \quad u(t) = \zeta_o + \zeta_1 \sum_{i=0}^{\infty} \eta^i \ln I(t - 1 - i)$$

where

$$0 \leq \eta < 1 .$$

It is well known that a distributed lag of this form can be expressed as a one-period lag.[2] Subtracting $\eta u(t - 1)$ from u(t), one finds that

$$(4.4) \quad u(t) = \zeta_o(1 - \eta) + \zeta u(t - 1) + \zeta_1 \ln I(t - 1) .$$

Next, substituting into (4.4) the value of u(t) given by (4.1), and simplifying,

$$(4.5) \quad \ln A(t) = \zeta_o' + \zeta_1' t + \zeta_2' \ln A(t - 1) + \zeta_3' \ln I(t - 1)$$

where

$$\zeta_o' = (1 - \eta)(\alpha + \zeta_o) + \lambda \eta$$

$$\zeta_1' = \lambda(1 - \eta)$$

$$\zeta_2' = \eta$$

$$\zeta_3' = \zeta_1 .$$

The parameters in equation (4.5) can be estimated and solved in a straightforward manner to obtain the original coefficients. This procedure enables estimation not only of the influence of the immediately preceding investment on the level of technology, but also of the rate of decrease of

---

1. L. M. Koyck, Distributed Lags and Investment Analysis (Amsterdam: North Holland Publishing Co., 1954). See also Marc Nerlove, Distributed Lags and Demand Analysis for Agricultural and Other Commodities, U.S. Department of Agriculture Handbook No. 141 (Washington: U.S. Government Printing Office, 1958).

2. See Koyck, op. cit., and Nerlove, op. cit.

this influence over time, assuming that this can be approximated by a decaying geometric series. Moreover, it offers the advantage of enabling one to relate the level of technology to the total amount of investment during an interval. [1]

It is not necessary to restrict oneself to a single explanatory variable, and the inclusion of others, if they add to the significance of the regression, may provide a more complete explanation of productivity change. One distinction which might prove useful is that between expenditure on structures and that on equipment. Fortunately, capital expenditure is broken down into these two categories (in addition to inventory change) by the U.S. Department of Commerce, and while there is undoubtedly some arbitrariness in this bifurcation, the underlying concepts seem relatively clear in most cases. In general, one would not expect the technology-improving effect of a dollar's worth of building to be equal to that of an equal amount of machinery; one would think the former would be decidedly inferior as a vehicle for improvements. Yet, although new techniques of production are usually associated with machinery rather than with structures, it is not obvious that the former will prove more strongly correlated with shifts in A(t). A firm may make routine, year-to-year improvements within existing facilities, incorporating the new designs in more modern types of equipment. Provided the modifications are of a minor nature, they will not, perhaps, necessitate vacating the plant. But the more drastic changes in the production function are likely to require scrapping the existing arrangement in its entirety and moving to new quarters within which the new processes can be installed. If this is the way in which innovations occur, [2] the more radical refinements in the state of the arts will be accompanied by expenditure on plant in addition to that on equipment, and the coefficient associated with the former may well exceed that associated with equipment expenditure.

Employing the finer classification of investment, equation (4.5) can be rewritten

$$(4.6)\quad \ell nA(t) = \zeta_o^* + \zeta_1^* t + \zeta_2^* \ell nA(t-1) + \zeta_3^* \ell nI_E(t-1) + \zeta_4^* \ell nI_s(t-1),$$

where I_E and I_s denote expenditures on equipment and structures, respectively, divided by the capital stock.

---

1. There is some possibility of an identification problem here. The fact that investment is lagged behind technical change reduces the likelihood of this difficulty, but does not eliminate it, particularly if there is serial correlation in either the investment or the technical change series.

2. This hypothesis was suggested to me by Professor Ingvar Svennilson (in a conversation).

Empirical Results

For the dependent variable, the Section II calculations of A(t) were used for the period 1919 to 1958. For the independent variables, expenditure by manufacturing firms on equipment and structures, respectively, for the years 1918 to 1957[1] were divided by the Section II capital stock figures, adjusted for idle capacity.

The estimates of the coefficients in equation (4.6) are presented below, with the t-ratios in parentheses, below their respective coefficients.

$$\ell nA(t) = \underset{}{.103} + \underset{(2.68)}{.068t} + \underset{(5.43)}{.631}\ \ell nA(t-1) + \underset{(.05)}{.005}\ \ell nI_E(t-1) - \underset{(-.82)}{.050}\ \ell nI_s(t-1)$$

The coefficients, ξ_3^* and ξ_4^*, can be interpreted as indicative of the effect on the level of technology of a hypothetically higher rate of expenditure on equipment and structures, respectively, at a given point in time. Neither of these coefficients is significantly different from zero at the .05 level.

Turning to the other estimates, it is readily seen that ξ_1^*, the coefficient of time, is significantly positive at the .01 level. The coefficient of $\ell n A(t-1)$, ξ_2^*, is also positive and significant at .01. Had ξ_2^* been negative, it would have indicated that the model was incorrectly formulated: the distribution of the lag must be positive.

The results do not lend support to the hypothesis that investment has acted as a vehicle for improvements in the technology. In view of this, I have estimated the coefficients in equation (4.5), where investment is not subdivided. The results are

$$\ell nA(t) = \underset{}{.123} + \underset{(2.78)}{.069t} + \underset{(5.82)}{.649}\ \ell nA(t-1) - \underset{(-.94)}{.058}\ \ell nI(t-1)$$

The estimated coefficients for t and $\ell nA(t-1)$ are not materially different here than in equation (4.6), either regarding value or statistical significance. Both coefficients are significantly greater than zero at the .01 level. The coefficient for investment again fails to be significant at the .05 level, however.

Interpretation of the Results

The preceding results could be advanced as evidence that there is no positive relationship between improvements in the technology and capital expenditure, and consequently that the induced component of technical change is non-existent. In this case, the role of investment in the growth process has been as unimportant as was suggested by the earlier studies. This is by no means a necessary conclusion to be drawn, and an alternative

1. Lowell J. Chawner, "Capital Expenditures for Manufacturing Plant and Equipment—1915 to 1940," Survey of Current Business, XXI, No. 3 (March 1941), p. 10; U.S. Department of Commerce, U.S. Income and Output, p. 216. Expenditures were deflated and expressed in 1954 dollars.

explanation, in fact, appears more reasonable. Consider equations (4.1)–(4.6). What do they really imply about the process of innovation? It might be well to recollect that the measure of technical advance discussed in Section II involves a subtraction of the first difference of the capital stock (duly weighted and expressed as a ratio of the total stock) from that of output for the same year, thereby implying that all of the net capital formation results in increased capacity immediately. To the extent that the expansion in capacity fails to materialize until subsequent periods, the data will indicate an increase in the capital stock unaccompanied by a corresponding rise in output (more precisely, an increased deepening without a rise in productivity). This results in an understatement of the shift in the production function for the period in question. This effect is likely to be highest when investment is greatest, so that any positive relationship between the two variables would be obscured. That is, a higher level of investment may lead to a larger understatement of the level of technology.

One way to test this newly proposed hypothesis is to regress $\frac{\Delta q}{q}$ against $\frac{\Delta k}{k}$. From equation (2.3a), by rearranging terms, one can write

$$(4.7) \quad \frac{\Delta q}{q} = \frac{\Delta A}{A} + \beta \frac{\Delta k}{k} \quad .$$

Differentiating (2.8) with respect to time results in

$$(4.8) \quad \frac{\Delta A}{A} = \lambda \quad .$$

Substituting (4.8) into (4.7),

$$(4.9) \quad \frac{\Delta q}{q} = \beta \frac{\Delta k}{k} + \lambda \quad .$$

A regression of $\frac{\Delta q}{q}$ against $\frac{\Delta k}{k}$ should result in independent estimates of λ and β, call them, respectively, λ^* and β^*. These estimates will be unbiased only if there is no induced technical change (so that MMP and AMP are equal). If, on the other hand, increases in the capital stock tend not only to enlarge capacity, but also to improve technology (so that MMP will exceed AMP), then the estimates will be biased. In this case, $\beta^* > \beta$ and $\lambda^* < \lambda$, because β^* will reflect that part of the increase in output attributable to induced technical change in addition to that part due to capital formation, whereas λ^* will represent only autonomous technical change. In any event, whether or not there is any induced innovation, the estimated slope coefficient should be significant due to the capacity-creating effect of capital formation.

Estimating the coefficients of (4.9) by least squares yielded the following result:

$$\frac{\Delta q}{q} = \underset{(4.17)}{.0325} - \underset{(-.57)}{.0416} \frac{\Delta k}{k}$$

That the slope parameter is not significantly positive supports the suspicion that the results of the earlier test need not be taken at their face value. Absence of a discernible positive relationship between increments in the capital stock per manhour and in productivity is certainly sufficient explanation for the failure of shifts in the production function to be associated with gross investment. This means that it is impossible to link specific productivity increases with specific instances of deepening. But is this more damaging than it appears at first sight? What implications does this fact have regarding the indices of technology constructed in Sections II and III?

I submit that the difficulties raised by these latest results have virtually no effect on the estimate of λ; they do have considerable effect on the annual values of $\frac{\Delta A}{A}$, but the trend is little affected. I have arbitrarily assumed that the productivity-increasing effects of investment would all occur within the period in which the expenditure is made. There are several reasons why this may not be so. First, there may be some building ahead of demand such that the capital is put into place long before it becomes operational,[1] creating a discrepancy between capital services and physical capital. But probably much more important, the capital data may partially reflect expenditures on plant and equipment which have not yet been completed, so that their effect on productivity during the first period that they are carried on the books is nil. But if this is the case, what meaning can be attached to a measure of technical change which assumes that all the benefits from an asset begin to accrue from the moment the asset is carried on the records? Too much confidence must not be placed in the annual figures, but if several years are averaged, the resulting figures most likely convey a true picture of the pattern of change. What has been done is to use as a proxy variable the following:

$$k^* = k + v ,$$

where there is a strong presumption that the v's are related by a first order autoregressive equation,

$$v_t = \rho v_{t-1} + \epsilon_t \qquad \rho < 0 ,$$

where ϵ is a stochastic element. The variable which is conceptually most appropriate to the model is k, but what we observe is k^*. In any one year, the two will, in general, differ in value by an amount, v. But the relationship among the v's is such that an understatement of the addition to the capital stock for a particular year is likely to be offset by a corresponding overstatement the following year. An average of the k^*'s can be expected to be quite close in value to that of the k's, for a period of several years.

Moreover, this same factor offers a plausible explanation for the

---

1. See H. B. Chenery, op. cit.

negative annual rates of technical change observed in Section II.[1] If, for the reasons discussed above, capital is carried on the books before it leads to an increase in capacity, there will be an addition to the capital stock without a corresponding increase in output. This will result in a decline in the level of technology, as estimated by the Section II model. This does not indicate a worsening of technical knowledge; rather, it suggests that the variables which have been observed are not the ones specified by the model.

An Interindustry Comparison

The conclusions emerging from this section are, thus far, principally negative in character. The hypothesized interrelationship between technical change and gross investment is not supported by the time series data. If induced technical change plays a major role in the innovative process, then it appears unlikely that time series analysis will shed much light on this fact because of the difficulty of obtaining appropriate capital input measurements. Available data do not permit associating increments in output with specific units of investment, and thus are too crude to be of value in investigating the hypothesized complementarity between investment and technical improvements.

Disaggregating to the SIC two-digit industry level, however, enables the use of cross section data instead of time series data. Rather than attempting to relate the aggregate level of technology to the age composition of the capital stock for each year, one can consider an interindustry comparison of technologies so as to by-pass the difficulties encountered above. Considering the level of technology in 1957 as the dependent variable, one can try to explain the interindustry variation in terms of differences among industries in the age composition of the capital stock, using as an index of the "newness" of the stock the ratio of gross investment to net capital stock, averaged over the period 1949-57.[2] Thus, according to the discussion

1. This is partly substantiated by the fact that years with negative technical change tend to be either preceded or followed by years with unusually large positive technical change. The notable exceptions to this are the years 1945 and 1946, during which period large structural changes occurred as a result of demobilization.

2. If the dependent variable were, instead, improvements in technology, there would be a serious identification problem. For, industries which are experiencing rapid autonomous technical change will be buoyant, and will provide a strong incentive to invest. Thus, a positive correlation between technical advance and investment could not readily be interpreted as evidence of induced technical change.

Even in the present case, however, problems of identification cannot be ruled out. In particular, if the level of technology is positively correlated with changes in technology (in the recent past, for example), then high-technology industries are likely to attract investment. However, there is no conclusive evidence to support this correlation.

above, one would expect the level of technology to be highest in those industries with a relatively new stock. The technology level of an industry is, of course, a result of many forces operative during the period; I am hypothesizing the age structure of the industry's capital goods to be one of these forces.

In order to compare industries with regard to the level of technology, it is necessary to impose an additional restriction on the form of the production function. Measures of the technology level have thus far been expressed as index numbers, enabling intertemporal comparisons within an industry but not interindustry comparisons at a point in time. This difficulty can be overcome by assuming that the production possibilities open to each industry can be characterized by a Cobb-Douglas function (though the parameters are not constrained to be the same in all industries). This additional assumption permits writing:

$$(4.10)\quad q_i(t) = A_i(t)\, k_i(t)^{\beta_i} \quad ,$$

where the subscript designates the industry. Solving (4.10) for A,

$$(4.11)\quad A_i(t) = \frac{q_i(t)}{k_i(t)^{\beta_i}} \quad .$$

$A(t_F)$ can be regarded as the dependent variable and $\bar{I}$ as an independent variable, where $\bar{I}$ is the ratio of gross investment to the capital stock, averaged over the period, and t_F refers to 1957, the final year of the period. If the hypothesized interdependence is present, the coefficient of $\bar{I}$ should be positive.

Several other independent variables can also be considered. One rather obvious one, in view of the preceding discussion, is the ratio of equipment to structures. While capital stock data are not broken down in this way for the industries, there are data relating to gross investment in each category. By summing gross investment in equipment and in structures, respectively, over the whole period, and then taking the ratio of the two, one can get a fairly good approximation to the ratio of the two stocks. In other words, I am defining a variable, $\bar{R}$, which is given by:

$$\bar{R} = \frac{\Sigma\, E(t)}{\Sigma\, S(t)} \quad ,$$

where E(t) and S(t) equal, respectively, gross investment in equipment and in structures, in year t. One might expect the efficiency level of an industry to be negatively correlated with $\bar{R}$, for the reasons given earlier.

A third independent variable to consider is the ratio of "fixed" capital to "variable" capital, although it is not clear a priori just how this might affect the efficiency level. This variable is designated $\bar{J}$, and refers to the average annual value of the fixed capital-inventory ratio, where fixed

capital is the sum of equipment and structures.

The final independent variable to be considered is labor "quality." One might venture a guess that this would be one of the most significant factors responsible for differences among industries with regard to the technology level. There is sufficient evidence of the heterogeneity of labor to force one to recognize the possibility of systematic interindustry differences in the types of skills and abilities required. This, in turn, would mean that a measure of technology based on the assumption of labor homogeneity would tend to overstate the technology level in industries using relatively high labor quality relative to those using less skilled workers. In other words, differences among industries in the quality of labor used will appear as differences in the technology level.

As a measure of labor quality, we can use the average wage rate for the period, $\overline{W}$. An industry which pays, on the average, a high wage is thus assumed to employ more highly skilled labor. Caution is required, however, in interpreting the coefficient associated with this variable, as the causation can obviously run either way. That is, a high level of efficiency in the industry can result in high wages just as well as high labor quality can lead to a high efficiency level.

The resulting regression equation can be written,

$$(4.12)\quad A = \alpha_0 + \alpha_1 \overline{I} + \alpha_2 \overline{W} + \alpha_3 \overline{J} + \alpha_4 \overline{R} \ .$$

Results of the Interindustry Analysis

To estimate the coefficients in equation (4.12), I used data for the 19 SIC two-digit manufacturing industries for the period 1949 to 1957. The estimates are presented below (with the t-ratios in parentheses below their respective coefficients):

$$\begin{array}{lllll} A = .20 + & .07\ \overline{I} + & .99\ \overline{W} - & .09\ \overline{J} - & .06\ \overline{R} \\ & (3.69) & (5.31) & (-2.51) & (-1.40) \end{array}$$

It can be seen that while α_4 is not significant at the .05 level, the other coefficients are all significant at the .01 level. α_3 is negative, signifying that the technology of an industry has tended to vary inversely with the fixed-variable capital ratio. As was expected, α_2 is positive, evidencing a direct relationship between labor quality (or, at any rate, the wage rate) and technology. Finally, it can be noted with pleasure that α_1 is significantly positive, indicating that the technology level is related to the age of the capital stock, and in the way initially hypothesized.

It is interesting to note the quantitative significance of the relationship between $\overline{I}$ and A. One measure of this is the partial elasticity of A with respect to $\overline{I}$, as calculated at the means of the two variables. This is given by:

$$\delta = \frac{\partial A}{\partial \overline{I}} \cdot \frac{\overline{I}}{A} = \alpha_1 \frac{\overline{I}}{A} \ ,$$

which is found to equal .25, indicating that a 1% rise in the investment-stock ratio results in a .25% increment in the level of technology.

Omitting $\bar{R}$, equation (4.12) can be rewritten:

$$(4.12)\quad A = \alpha_0' + \alpha_1' \bar{I} + \alpha_2' \bar{W} + \alpha_3' \bar{J} \quad .$$

The estimates are:

$$A = -.02 + \underset{(3.81)}{.07}\ \bar{I} + \underset{(5.44)}{1.03}\ \bar{W} - \underset{(-3.79)}{0.12}\ \bar{J} \quad ,$$

which do not differ appreciably from the corresponding estimates in (4.12).

Finally, it can be noted that if $\bar{J}$ is dropped, $\bar{R}$ becomes significant. Write:

$$(4.14)\quad A = \alpha_0^2 + \alpha_1^2 \bar{I} + \alpha_2^2 \bar{W} + \alpha_4^2 \bar{R} \quad .$$

This set of estimates is given by:

$$A = .36 + \underset{(2.49)}{.05}\ \bar{I} + \underset{(4.51)}{.98}\ \bar{W} - \underset{(-2.85)}{.12}\ \bar{R} \quad .$$

All coefficients are significant at .01, and $\bar{R}$ now turns out to be negative, as suggested by the earlier discussion.

The results of the cross section study are fully compatible with the hypothesis that technical change is closely related to the age structure of the capital stock, and hence to the level of investment in the recent past. While the time series analysis did not support this conclusion, neither were the data incompatible with the conclusion if one accepts the test on pages 341–42. This test demonstrates the inadequacy of the time series data with regard to the primary question asked in this section. Thus, one may tentatively conclude that there is some evidence, on the basis of an inter-industry comparison, that advances in technology are partially a result of investment activity. This, in turn, suggests the need to qualify the results of Section II, which suggest investment to be relatively insignificant as a contributory factor in the growth process.

Additionally, the present results support the hypothesis that large changes in technology are accompanied by expenditure on structures, though it is obvious that alternative interpretations would also be consistent with the data. Moreover, technology seems to be inversely related to the ratio of fixed capital to inventory. Finally, a positive relationship was noted between the technology level and the wage rate; while the latter was used here as an index of labor quality, one must be cautious in giving this interpretation. In particular, the positive relationship can easily be a result of a tendency for wages to be high in technologically advanced industries.

5. Conclusion

The preceding analysis has been organized primarily around three

dichotomies. (1) Capital deepening and technical advance have been suggested as the two agents responsible for productivity change, and estimates of their relative significance were presented. (2) Technical change was in turn broken down into an interindustry and an intraindustry component, and the observed improvements in technology were allocated between the two. (3) Finally, the distinction was drawn between autonomous and induced shifts in the production function, and an attempt was made to gauge their relative importance in the manufacturing sector. Now we turn to summarize the empirical results of the study and to discuss what conclusions one may draw.

Despite the observed interrelationship between technical change and investment, one can conceptually distinguish between the two. On the basis of this distinction, the present study has found technical change far more important than capital deepening in bringing about a rise in output per manhour during the past four decades. The present essay, then, confirms the findings of other scholars on this point.[1]

This is not to attribute any causal significance to the results presented here. From the methods used, it is not possible to say whether the historical relative impact of technical change has been great because (1) technical change tends to promote increases in output more readily than does investment, or because (2) we have had more technical change than capital deepening during the period in question. In particular, one cannot use the conclusions of this essay, without further evidence, to support the view that more resources should be devoted to improving the technology and fewer resources to increasing the rate of investment. To make this assertion, one would have to be able to demonstrate that the marginal productivity of a dollar spent on research and development exceeds that of a dollar spent on expanding the capital stock—an interesting hypothesis but not one with which the present paper has been concerned.

Section III of the present study indicates that the rate of technical change in aggregate manufacturing has exceeded a weighted average of the industry rates of technical change by a substantial margin. The discrepancy between the two, which we have termed interindustry technical change, amounts to nearly one-third of the aggregate manufacturing rate of technical change. Most of this interindustry technical change can be attributed to shifts of capital from industries where its marginal productivity is low to those where it is higher.

That interindustry technical change is so great is not surprising in view of the observed dispersion among industries in the estimated rates of technical change. This dispersion provides a large potential rate of interindustry technical change. To take advantage of this potential, it is clearly important to maintain interindustry factor mobility.

Finally, in Section IV, the interrelationship between investment and improvements in technology was examined. The cross section data showed

---

1. See footnote 1 on p. 304.

that industries with high levels of technology tended to have both high capital turnover ratios and high wage rates. While, again, no causal significance should be attributed to these results, the relationship between technology and capital turnover ratios is consistent with the hypothesis that, as one would expect, investment has tended to act as a vehicle for bringing about technical change.

Summary: This paper is concerned with one aspect of the dependence of Canadian monetary policy upon that of the United States, namely, the relationships between the capital markets of the two countries. The theoretical analysis is supplemented by an empirical investigation of the period 1952–1960, during which time there existed complete convertibility of both U.S. and Canadian dollars and, in addition, a willingness on the part of the monetary authorities of both countries to pursue flexible interest rate policies.

With the exception of the provincial governments, it cannot be established that exchange rate expectations are of significant importance to Canadian borrowers. It is found that the largest component of the long-term capital movements between the two countries is influenced by the interest rate differentials between them, and this is concluded to be a major contributing factor to the close links which are found to exist between equivalent U.S. and Canadian interest rates. The results show that all types of Canadian issuers of bonds and debentures have had a significant part in this tying together of rates. However, since new issues occur only at irregular intervals, the trade in outstanding securities must take the major responsibility for maintaining the close relationships between the capital markets. This appears to occur mainly through the trade in Canadian provincial, municipal, and corporate securities and in U.S. Government bonds.

This essay was submitted to Yale University in 1962 in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The author is currently Assistant Professor of Economics at Yale.

Dissertation Committee: Professors Robert Triffin (Chairman), Michael C. Lovell, and Raymond P. Powell.

*

I am grateful to Professors Michael C. Lovell, Raymond Powell, and Robert Triffin for comments upon an earlier draft; to E. B. Carty of the Dominion Bureau of Statistics and representatives of A. E. Ames and Company, Moody's Investors' Service, McLeod, Young, Weir and Company, and Moss, Lawson, and Company, for assistance in the collection of statistics; and to the Ford Foundation's Doctoral Dissertation Fellowship Program and the Cowles Foundation for Research in Economics at Yale University for financial assistance.

G.K.H.

CONNECTIONS BETWEEN UNITED STATES' AND CANADIAN CAPITAL MARKETS, 1952-1960

Gerald K. Helleiner

Discussion of Canadian monetary policy invariably refers to its dependence upon concurrent monetary policy in the United States—a dependence which, it is generally believed, has diminished since the introduction of a flexible exchange rate in 1950. The role which ties between U.S. and Canadian capital markets play in this dependence is, obviously, a crucial one. It is the intention of this essay to present theoretical and empirical evidence concerning the extent and nature of these ties. The empirical analysis concerns a period (1952-1960) during which there existed complete convertibility of both U.S. and Canadian dollars and willingness in both countries to pursue flexible interest rate policies.

Parts A and B of this essay, which concern interest rate connections and international portfolio capital movements, respectively, contain the bulk of this analysis. Sections I, II, and III of Part A discuss, in theoretical terms, the nature of the ties between U.S. and Canadian long-term equivalent interest rates. Sections IV and V offer empirical results of analysis of these interest rates during the 1952-1960 period. In Part B, the results of regression analysis conducted on U.S.-Canadian portfolio capital movements are presented. Section I discusses the independent variables used, and sections II and III contain, respectively, analyses of new issues and the trade in outstanding issues. Part C very briefly relates the earlier analysis to the question of the dependence of Canadian monetary policy upon that in the U.S.

A. Connections between United States' and Canadian Long-Term Interest Rates

I. Unifying Influences

The ability of capital to flow freely across international borders without restriction or, more properly, its ability to move freely in and out of different currencies is the most important factor tending to bind separate capital markets together. Where this freedom exists, investors may seek out the highest yields and borrowers the lowest possible lending costs without regard for international boundaries. In doing so, of course, they tend to equalize interest rates in all markets. Since interest rates are "pure" numbers, this equalization occurs regardless of the rate of exchange between currencies at any particular time.

At first glance it might seem that ties between Canadian and United States' long-term interest rates result solely from the international flow of long-term securities themselves. This need not, however, be the case. Indeed, Morgenstern's results from investigation of the pre-World War I European experience suggest that the major explanation for the existing ties between long-term rates was the extremely close connection between short-term rates.[1] Short-term rates are bound to one another through flows of

1. Oskar Morgenstern, International Financial Transactions and Business Cycles (Princeton: Princeton University Press, 1959), Ch. IX.

short-term capital, which are notoriously responsive to interest rate differentials. Since long-term rates are related to short-term rates within each separate economy, it follows that there may exist strong ties between capital markets, even where there are no long-term international capital flows at all.

Such indirect relationships between long-term interest rates may be far from perfect. In the first place, the strength of the domestic relationship between short-and long-term rates may vary from country to country or from period to period. In the second place, even the direct ties between short-term rates may be somewhat imperfect. This possibility is heightened where there exists a forward market for foreign exchange which permits sophisticated investors to be content even though short-term interest rates are divergent. [1]

International capital flows of still another type could also conceivably result in ties between long-term interest rates in the United States and Canada. To the extent that flows of direct investment respond to international differentials on the yield of real assets (marginal productivity of capital), there will be a tendency towards the international equalization of these yields. Since in both countries yields on financial assets are not completely unrelated to those on real assets, these flows will also tend to equalize long-term interest rates. But it is the large flow of long-term securities between the United States and Canada which remains by far the most likely influence on the degree of interconnectedness of their two capital markets.

However, capital flows, of whatever type, are not a prerequisite for the existence of close interrelationships between capital markets. The presence of a potential capital flow may be quite sufficient to keep interest rates in separate capital markets together or moving sympathetically. Close working relationships between Canadian and U.S. stock exchanges, dealers, and businessmen of all types undoubtedly make this point of considerable importance.

II. Factors Permitting Divergent Behavior of Interest Rates

The above-mentioned unifying influences do not prevent divergences in levels and movements of equivalent U.S. and Canadian interest rates; rather, they merely set certain limits upon them. For analytical purposes, it may be useful to consider the various factors other than international interest rate differentials which may affect the international investment and borrowing decisions of U.S. and Canadian capital market participants. If such other factors are sufficiently important, the unifying tendencies may be completely counteracted.

Analysis of these other factors may be based upon either of two theoretical approaches to the behavior of participants in capital markets. The

1. Equilibrium is possible when the forward discount on one currency exactly offsets the higher rate of return which could be earned by holding that particular currency.

first postulates that the market participant bases his decisions as to whether to invest (or borrow) at home or abroad upon the value of only one variable—the international differences between expected returns (or costs). The international interest rate differential is obviously a major component of the international difference between expected returns (costs), but there exist others. In this formulation, risk considerations enter only in the form of premia or discounts on the expected returns (costs). A market participant behaving in this manner, it is generally assumed, will always make all-or-nothing decisions. Depending on the international difference between expected returns (costs), he will either invest (borrow) his entire portfolio in domestic assets (liabilities) or in foreign ones; there will be no domestic-foreign security mix. If changes in the international difference between expected returns affect capital flows, they do so by stimulating individual market participants on the domestic-foreign margin to make 100 per cent portfolio readjustments.

The second approach postulates that the individual market participant bases his decisions not only upon considerations of expected return, but also upon those of risk, to which he is, in general, assumed to be averse. It has been demonstrated that such behavioral assumptions permit the investor (or borrower) to hold diversified portfolios of assets (liabilities).[1] Changes in the international difference between expected returns now result in capital flows through their stimulation of all individual market participants to make partial readjustments in the domestic-foreign composition of their portfolios. Since portfolio diversification with respect to currency of repayment is characteristic of both borrowers and investors in the international capital market, this approach appears more relevant. The first approach described above is not, however, totally incompatible with the existence of portfolio diversification once certain quite realistic qualifications are introduced. With investor (borrower) realization that too heavy sales or purchases will destroy a relatively thin market's advantages, the existence of a number of separate time horizons each of which is relevant to a different component of the individual's portfolio, the inability of borrowers to recall outstanding issues, or the existence of a degree of sheer investor inertia, portfolio diversification and partial readjustment may exist even where expected return is the only decision variable.

Whichever approach one adopts, the factors influencing decisions in the international capital market are the same. Changes in any of them will produce international capital flows in the same direction, regardless of which of the assumed underlying behavior characteristics is more realistic.

---

1. See, for example, Harry M. Markowitz, Portfolio Selection, Cowles Foundation Monograph 16 (New York: John Wiley and Sons, 1959); J. Tobin, "Liquidity Preference as Behavior Towards Risk," Review of Economic Studies, XXV (2) (February 1958), pp. 65–86; Evsey D. Domar and Richard A. Musgrave, "Proportional Income Taxation and Risk-Taking," Quarterly Journal of Economics, LVIII (May 1944), pp. 388–422.

These factors may be grouped under two headings: 1) expected returns, under which are included exchange rate expectations, divergences in interest rate expectations, transactions costs, tax considerations, and thinness of markets; and 2) risk, which includes exchange risk, consistent differences in security characteristics (default risk), and risk reduction through portfolio diversification. Their existence, along with a third group of factors which may be labelled adjustment imperfections, permit divergences in the behavior of equivalent long-term interest rates in separate capital markets, even when there is a continuously free and elastic flow of capital between the two markets.

1. Expected Returns

(i) Exchange Rate Expectations. The expectation of an appreciation in the value of a currency will obviously render investments the interest and/or principal of which are payable in that currency more attractive. Thus, in order to induce investors to undertake such investments or refrain from disposing of them, a smaller nominal interest rate advantage is necessary. Similarly, borrowers will be less willing to incur obligations payable in the currency which they expect to appreciate and will only be persuaded to do so by a larger nominal interest rate advantage.

The subjective exchange rate expectations of an individual participant in the international capital market may be described by a probability distribution. Such a distribution might show the probability which the individual ascribes to the realization of all possible rates of exchange between the United States and Canadian dollars at some future date. Of greater analytical interest, however, would be another distribution which would show, instead of expected future rates of exchange, the individual's expectations as to the possibility of capital gains or losses, expressed as rates per annum, resulting from changes in the exchange rate between the present and some fixed future date. It has become conventional practice in portfolio selection theory to employ the expected value of such probability distributions as this one as a summary variable for expectations. In the interest of clarity this procedure is employed here.

Let us call the Canadian (U.S.) individual's expected value of the per annum rate of capital gain (loss) from holding United States rather than Canadian (Canadian rather than United States) dollars g_1. If the U.S. dollar is expected to depreciate relative to the Canadian dollar g_1 will, of course, be negative. Other things being equal, participants in the capital market will be indifferent as to whether to buy (or sell) U.S. or Canadian dollar securities when the Canadian interest rate (R_c) less the equivalent U.S. interest rate (R_u) equals g_1. Exchange rate expectations, for which g_1 is a summary, are obviously a major determinant of market participants' decisions. These expectations are, of course, continually subject to change.

The time horizon of the individual in question is of some importance to the size of g_1. An investor in short-term instruments need only concern himself with the behavior of the exchange rate in the near future. The same is true of the investor in long-term securities who intends to dispose of them

quickly. On the other hand, the long-term investor or the issuer of bonds need only be concerned with the level of the exchange rate in the more distant future.

When the prospective capital gains of short-term investors are expressed as yields per annum, they are likely to exceed in size those of longer-term investors. An expected rise in the value of the Canadian dollar merely from $1.000 in one month to $1.005 in the next, for example, implies a capital gain or loss of 6 per cent per annum for holders of foreign securities. Given a reasonable degree of long-run exchange rate stability, the exchange rate is not likely to be of crucial importance to long-term investors or borrowers although it will still pay them to choose carefully the times for issues, redemptions, sales, and purchases. As Keynes pointed out, if the exchange rate is subject to a 5 per cent total fluctuation on either side of par, "then a 5 per cent loan in terms of the lender's money remitted at the par of exchange may cost in future years anything between 4-3/4 per cent and 5-1/4 per cent interest, and when it is paid off the redemption may cost anything between 95 and 105, in terms of the borrower's money. In the case of a long-period loan these possibilities are not very serious. . . ."[1] Assuming random fluctuations about par, the extra-large interest coupons resulting from an unfavorable (from the borrower's viewpoint) exchange rate will approximately offset the extra-small ones resulting from a favorable rate. The major risk then only concerns the redemption date and this can be all but eliminated, as far as the borrower is concerned, through sufficiently liberal call privileges. The possibility of irreversible trends in the rate of exchange, of course, remains. It appears from this discussion, then, that g_1 is likely to fall as the relevant individual's time horizon is extended.

(ii) Divergences in Interest Rate Expectations

Capital gains other than those created by exchange rate fluctuations are also relevant. These may arise domestically as a result of changes in security prices which are related to interest rate changes. On an international level, expected returns may differ between countries as a result of different interest rate expectations. Let us call the individual's expected value of the per annum rate of capital gain (loss) from holding U.S. rather than Canadian (Canadian rather than U.S.) dollars g_2. If g_2 is positive, the individual's implied expectation is that U.S. interest rates will rise during the relevant time period to a greater extent than their Canadian counterparts. Like g_1, g_2 is subject to continual change and is likely to fall as the relevant individual's time horizon is extended.

(iii) Transactions Costs. The major element of international

1. John Maynard Keynes, Treatise on Money (London: Macmillan and Co., Ltd., 1930), Volume 2, pp. 333-34. See also C. G. Bale, "On the Exchange Risk Involved in Borrowing Abroad," Canadian Journal of Economic and Political Science, XXVII (February 1961), pp. 97-100.

transactions costs is the commission which must be paid to foreign exchange dealers upon conversion of the domestic currency to foreign currency and back again. This commission is customarily stated as a percentage of the amount being exchanged. For the purposes of this analysis, this cost must instead be expressed as a rate per annum spread over the length of the time period of relevance to the individual in question. This rate, which will be called c, diminishes as the time horizon of the individual lengthens, given the size of the transaction involved.

Extra transactions costs may arise if such costs as underwriters' and brokers' commissions, legal fees, and expenses incurred in the pursuit of information are higher in the one capital market than in the other. If this were the case, the effect upon market participants' decisions would be asymmetrical, depending upon whether the individual in question were a Canadian or a U.S. resident. There exists little statistical evidence upon which to hang any statement concerning the relative costs of flotation or security sales and purchases in the U.S. and Canada.[1] Attention will therefore be confined to the question of exchange commissions, as represented by c.

Reintroducing exchange rate and interest rate expectations and holding other things equal, the international difference in expected returns is given by one of the expressions, $R_c-R_u-g_1-g_2-c$ or $R_c-R_u-g_1-g_2+c$, the relevant one depending on whether one is a U.S. or Canadian, an investor or a borrower. If one were to assume that borrowers and investors are "100 per cent readjusters," it would follow that a Canadian borrower would not issue U.S. dollar securities unless $R_c-R_u>g_1+g_2+c$; a U.S. investor would not invest in Canadian dollar securities unless $R_c-R_u> g_1+g_2+c$; and a Canadian investor would not invest in U.S. dollar securities unless $R_c-R_u< g_1+g_2-c$.

(iv) Tax Considerations. Relevant to a U.S. investor contemplating the purchase of Canadian securities or to a Canadian investor contemplating the purchase of U.S. securities are the tax rate he must pay to his own government, the tax rate he must pay to the foreign government, and the deductions allowed by his own government in recognition of his tax payments to the foreign government. His domestic tax rate will depend upon his total income and his tax status. The tax rate payable to the foreign government will depend, in the Canadian case, upon the type of security held. Both the U.S. and Canada prevent double taxation by allowing the deduction in full from domestic tax liabilities of taxes paid to foreign governments on portfolio investment income.

Without going into the details of the legislation, it is possible briefly

1. The New York market is apparently, however, a more expensive one in which to borrow than the London market. See E. Nevin, "Some Reflections on the New York Issue Market," Oxford Economic Papers (New Series), XIII (February 1961), pp. 84–102.

to summarize Canadian tax laws in this respect during the 1952–1960 period. The following withholding taxes on interest and dividends paid by Canadians to foreigners were called for: nothing on the interest paid on foreign currency loans, nothing on interest on Government of Canada bonds, 5 per cent on the interest from Provincial and Municipal (direct and guaranteed) bonds, and 15 per cent on the interest and most dividends from Canadian corporations' bonds and stocks.[1] Since new issues payable in U.S. dollars comprise the greater part of Canadian sales of securities to the U.S., tax questions do not, for the most part, even arise. The non-zero withholding taxes are sufficiently low that in most cases the U.S. investor will not suffer any increase in tax liability through his investment in Canadian rather than U.S. securities. The same may be said with respect to the Canadian investor in U.S. securities whose usual withholding tax payment to the U.S. Government on interest or dividends is at the rate of 15 per cent. There may, however, exist exceptions. U.S. tax-exempt institutions, for example, will be unwilling to invest in Canadian dollar securities unless the favorable interest rate differential is great enough to outweigh, among other things, the tax liability to the Canadian Government which they would thereby incur.

(v) Thinness of Markets. It is important to realize that the R_c and R_u in the above expressions should be the market interest rates which could be expected to prevail after the transactions being considered by the individual have been made. If, as is likely to be the case in Canada, the capital market is a thin (imperfect) one, the market participant must consider the effects of his own actions upon it. The response of the market to large transactions within it is, of course, quite variable. Less consideration need be given to this factor in the considerably deeper U.S. capital market. In principle, however, one should be using the difference between expected interest rates in the two markets assuming that the transactions had actually been made in both. Thus observed interest rates may not be the relevant measures of R_c and R_u.

2. Risk

(i) Exchange Risk. The possibility of changes in the rate of exchange between the U.S. and Canadian currencies exerts influence not only through the effects of expectations (g_1), but also through the effect of the mere existence of risk upon individuals' responses to international interest rate differentials. The possibility of a mistaken estimation of the future exists even when expectations are completely neutral (g_1 equals zero). An investor with an aversion to risk will avoid investments payable in a foreign

1. For more detailed discussion of Canadian tax laws relating to foreign investment see J. Grant Glassco, Royal Commission on Canada's Economic Prospects, Certain Aspects of Taxation Relating to Investment in Canada by Non-residents (Ottawa: Queen's Printer, 1956); and J. B. de la Giroday, Canadian Taxation and Foreign Investment (Toronto: Canadian Tax Foundation, 1955).

currency unless he receives compensation, in the form of a higher yield, for the exchange risk involved. (Risk-loving investors would, of course, be willing to pay premiums in the form of the acceptance of lower yields for the privilege of investing in foreign securities.) Analogous considerations influence the decisions of borrowers as to the currency in which they are to incur obligations.

The treatment of the exchange risk factor depends upon the assumptions made as to the nature of market participants' behavior. If expected returns are assumed to be the sole criterion for behavior (and all adjustments are 100 per cent ones), it is best to employ the concept of an exchange risk premium. This risk premium need not remain constant over time. Changes either in the individual's estimate of risk or in his degree of risk aversion may alter the compensation he requires. It is also probable that the necessary premium, expressed as a rate per annum, falls as the time horizon of the individual lengthens.

Let us call the premium which is required to compensate a risk-averting market participant for the bearing of exchange risk, expressed as a rate per annum, x. It follows that Canadian borrowers and U.S. investors will enter U.S. and Canadian capital markets respectively only when $R_c - R_u > g_1 + g_2 + c + x$; and Canadian investors will enter the U.S. market only when $R_c - R_u < g_1 + g_2 - c - x$.[1]

Using the alternative approach, which assumes that the market participant bases his decisions upon two variables, (1) the international difference between expected returns and (2) risk, as given customarily by some measure of dispersion (such as the standard deviation) of the distribution of expected returns from his portfolio, exchange risk plays a slightly different role. Exchange risk may be considered through the use of a measure of dispersion of the same distribution of which g_1 is the expected value. Decisions as to the domestic-foreign composition of assets or liabilities may involve a trade-off between expected returns and exchange risk.[2] Rather

---

1. Moody's Investors' Service implicitly estimates x in connection with the ratings it assigns to Canadian dollar securities. By convention, bonds payable in Canadian dollars are generally rated at one level below that at which they would have been rated had they been payable in U.S. dollars. The exchange risk premium which ought to be demanded by the U.S. investor is therefore postulated to be the difference between yields of the same class of bonds but of adjoining default risk ratings. This difference varies with the class of bond, the original rating of the bond, and the point of time being considered. This procedure is, in this respect, of course, nothing more than an arbitrary, if convenient, rule of thumb.

2. In some instances the incurring of exchange risk will not increase the expected return. No trade-off is then necessary. Both considerations will lead one to remain within the domestic market.

than making all-or-nothing decisions, as in the former approach, the market participant will alter the composition of his portfolio of assets or liabilities in response to changes in his estimates of exchange risk and expected returns. For example, when exchange risk is believed to rise, the international difference (in favor of the foreign market) in expected returns remaining unchanged, the market participant will reduce the percentage of foreign securities in his portfolio.

(ii) Consistent Difference in Security Characteristics (Default Risk). Strictly speaking, this factor should not be considered at all since the interest rates being compared internationally are supposed to refer to securities which are equivalent. It may be, however, that all Canadian securities differ from all U.S. securities, in which case true equivalents cannot be found. The most striking indication that such a situation may exist, at least in the minds of U.S. investors, is found in connection with Moody's Investors' Service bond ratings (of default risk), which undoubtedly guide or influence the decisions of large numbers of U.S. investors. Moody's rates the U.S. dollar obligations of the Government of Canada at "Aa"—that is, below the "Aaa" level accorded the obligations issued by scores of U.S. states, municipalities, and corporations. It appears, then, that even the safest issues of Canadian borrowers, issues which are not subject to exchange risk, are considered to be more risky (in terms of the likelihood of default) than the issues of many U.S. borrowers whose credit standing would not, on the face of it, approach that of the Canadian Government. This may be the result of the somewhat remote possibility of a severe exchange crisis in Canada leading to a moratorium on foreign interest payments.[1] Whatever the cause, if such a belief on the part of U.S. investors exists, it must either be compensated by an extra default risk premium on all Canadian securities or be considered in the calculation of the measure of dispersion which reflects the risk involved in Canadian securities. Assuming, as seems most realistic, that this feeling about foreign securities is asymmetrical—that is, Canadian investors do not consider U.S. borrowers more likely to default than equivalent Canadian ones—it need only be considered for the case of U.S. investors contemplating investment in Canadian securities. Using the "100 per cent readjustment" theory, U.S. investors will purchase Canadian securities only when $R_c - R_u > g_1 + g_2 + x + c + d$, where d is the default risk premium, expressed as a rate per annum which must be paid to U.S. investors because of their belief in the increased likelihood of default by Canadian borrowers. Since Canadian issuers of securities in the

1. Such beliefs could also, though not in this case, be the result of internationally divergent interest rate (capital value) risks. The latter risks bear the same relationship to interest rate expectations as exchange risk bears to exchange rate expectations. This possibility is mentioned here merely for the sake of completeness.

U.S. market will face the same U.S. investor belief, they will be forced to offer a higher interest rate than the going U.S. market rate. Hence this inequality is also the condition required for Canadian security issuers' entry into the U.S. capital market.

(iii) Risk Reduction Through Portfolio Diversification. Offsetting the exchange risk incurred by a capital market participant when he invests or borrows in the foreign capital market is the possibility of a reduction in his over-all risk, in particular his interest rate or capital value risk, as a result of the increased diversification of his assets or liabilities. The advantages to be expected from not "putting all one's eggs in one basket" depend upon the existence of a certain degree of independence between the expected experiences of the two "baskets." It will be seen that there exists a high degree of correlation between equivalent U.S. and Canadian interest rates. Advantages in the form of risk reduction to be gained from U.S.-Canadian portfolio diversification are not therefore likely to be great. To the extent that it may be considered important, however, this factor can be treated as a reduction either in the exchange risk premium or in the measure of dispersion being employed as the proxy for risk.

3. Imperfections of Adjustments

The preceding analysis implicitly assumed that the market mechanism operates relatively smoothly. There may, however, exist imperfections which affect its applicability to reality. For example, participants may be ignorant of the foreign opportunities open to them, they may be unwilling for non-economic reasons to borrow from or lend to "foreigners," or their responses to overseas capital markets may outweigh those to North American ones.

It is also worth mentioning that the smooth adjustment of interest rates may encounter difficulties with respect to the speed with which it takes place. Such adjustments are certainly not instantaneous: preparation for a bond flotation, for example, may take several weeks.

III. Factors Making for Divergent Average Levels of Interest Rates

It has been shown that, even in a smoothly working international capital market, there is no necessity for equivalent long-term interest rates to move together. A variety of factors which may prevent this have been discussed above. Not only is it quite possible for equivalent U.S. and Canadian interest rates to diverge in their movements, but they may also diverge in their average levels over long periods. There are a number of possible explanations for this.

(i) Exchange Rate Expectations or Divergences in Interest Rate Expectations. It is clear from the preceding discussion that such divergences in average levels would be possible through persistently one-sided exchange rate expectations (g_1) and divergences in interest rate expectations (g_2) by the bulk of market participants. Persistently positive or negative

evaluations of g_1 and g_2 by the transactors in the market are likely to produce a non-zero average for $R_c - R_u$.

(ii) Exchange Risk and Transactions Costs. The U.S. capital market is far larger than the Canadian capital market. The volume of U.S. funds available for investment in Canadian securities far exceeds the volume of Canadian funds available for investment in U.S. securities. Presumably, then, the aggregate response in terms of dollars of U.S. investors to changing Canadian opportunities is more elastic than is the aggregate Canadian response to U.S. opportunities. It follows that if the North American capital market is a unified one, exchange risk and transactions cost are much more likely to result in an extra premium on Canadian dollar securities than on U.S. dollar securities. Ceteris paribus, if exchange risk or transactions costs exert any influence at all, they will tend to result in Canadian interest rates which are higher than their U.S. equivalents.

(iii) Tax Considerations. It is obvious that where tax treatment of the income from domestic securities which are equivalent in every other respect differs, their before-tax yields will differ. This is also true where equivalent U.S. and Canadian securities are concerned, although the fact that investors in foreign securities are taxed by their own governments on their foreign income means that the incentives for international capital flows are left unchanged by tax considerations. Purely domestic pressures deriving from the differences in tax structure may be expected to affect relative interest rates. Thus tax-exempt municipal and state bonds in the U.S. should offer lower before-tax yields than their taxable Canadian equivalents. Similarly, since Canadian residents receive a large tax credit on dividend income and pay no capital gains tax, one would expect, other things being equal, that the before-tax yield on Canadian stocks would fall short of that on U.S. stocks.

(iv) Marginal Productivity of Capital. One of the commonest explanations of persistent international differences in the level of interest rates is the existence of differences in the marginal productivity of capital from country to country. Since investors in the high marginal productivity country have the alternative of investing in real capital, portfolio investment opportunities must offer them correspondingly higher yields if they are to attract any funds.

On a priori grounds, one would expect that capital's marginal productivity would be higher in Canada than in the U.S. because of Canada's vast unexploited natural resources and its lower capital per capita. Statistical evidence for this is lacking. However, the direct investment which has poured into Canada from the U.S. in unprecedented amounts during the 1950's suggests that, whether or not capital actually yielded a higher return in Canada, it was believed by many U.S. investors to do so, and, from the analysis' point of view, this is just as important.

(v) Government Policy. To a considerable extent, one would expect the over-all level of interest rates on any capital market to be determined by the monetary authorities. Thus, if, during the period, Canada had consistently used monetary policy more vigorously in an anti-inflationary direction than the U.S., one would expect to see it reflected in higher average Canadian interest rates than their U.S. equivalents. It is also possible that, as a result of different institutions or industrial structure, the interest rate implications of monetary policies with identical aims differ systematically between the two countries. Individual types of interest rates may also be affected by conscious monetary policy; for example, the term structure of rates may be influenced by the character of open-market operations although this possibility was restricted in the U.S. by the "bills only" policy in the 1950's. Non-monetary aspects of government policy may affect individual interest rates, too, as does the U.S. statutory limit on the interest rate to be paid upon long-term U.S. Government bonds.

(vi) Consistent Differences in Security Characteristics. The grounds for believing that U.S. investors may require compensation for a belief that Canadian securities are more liable to default than U.S. securities were stated in the previous section. If unity of the U.S. and Canadian capital markets exists, this factor would tend to raise Canadian interest rates relative to their U.S. equivalents.

(vii) "Structural" Differences. Average levels of equivalent interest rates may diverge for a variety of other reasons which may be summarized by the term "'structural' differences." These include such factors as the effects of different laws or customs regarding institutional investors' portfolios on the demand for particular types of securities, and the reputed cautiousness of the individual Canadian investor.

IV. Correlations between U.S. and Canadian Interest Rates, 1952–1960

There are two main questions which will be investigated in the following two sections. The first section examines the degree to which Canadian and U.S. interest rates on equivalent securities and the structures of these rates moved together during the 1952–1960 period. The second, much shorter, one explores the existence of long-run differences in the absolute level of rates on equivalent securities during the same period. Both matters have been discussed from a theoretical point of view in the preceding sections.

In order to meaningfully compare interest rates in different capital markets, it is necessary, first of all, to obtain statistics which pertain to securities which are as nearly equivalent as possible. An attempt has therefore been made to derive monthly estimates of U.S. and Canadian interest rates which are roughly comparable. Particular attention has been given to the degree of default risk, the term to maturity, and the type of borrower.

All of the time series employ monthly data which refer to the last

week of each month, although the exact day of the week is not the same for every type of security or even, sometimes, for the same security in successive months. Government of Canada bond yields are taken from the Bank of Canada's Statistical Summary. From the end of 1957 onwards, during which time the Bank of Canada no longer published Government of Canada bond interest rates by term to maturity, representative bonds of appropriate term to maturity were selected in order to continue these series.[1] The corresponding U.S. figures—interest rates on U.S. Government bonds of various terms to maturity—were taken from the figures published weekly by Moody's Investors' Service in its Bond Survey.

The method of deriving the yields on provincial (or state), municipal, industrial, and public utility bonds was the most troublesome of all. McLeod, Young, Weir and Company of Toronto compiles monthly Canadian bond yield statistics for each of these types of borrowers. Moody's does the same in the United States. In order to obtain default risk comparability, Moody's (default risk) rating of each rated Canadian bond used in the compilation of each type of McLeod, Young, Weir series was obtained.[2] Then allowing for Moody's evaluation of exchange risk and its apparent discounting of all Canadian-borrower securities, the Moody's yield series was selected for which the rating corresponds most nearly to the average ratings of the bonds in the equivalent Canadian series. Thus, because of corresponding differences in the ratings assigned to bonds in the various McLeod, Young, Weir series, the U.S. industrial and state bond yield series selected were those relating to Moody's "Aa" ratings, whereas the selected municipal and public utitlity series related to Moody's "A" ratings.

A somewhat similar procedure was employed in an attempt to achieve comparability of term to maturity. This time, however, in addition to a consideration of the component bonds of each Canadian series, it was also necessary to look at the component bonds of each Moody's series in order to arrive at "average" terms to maturity for each series. Each of the Canadian bond yield series was then adjusted in each month to eliminate the U.S.-Canadian difference in "average" term to maturity for that particular type of security. The method employed, admittedly rather a rough one, was to add to (or subtract from) the equivalent bond yield figure the difference between Government of Canada bond yields of terms to maturity which corresponded to the two "average" terms to maturity of the component

---

1. The Bank of Canada stopped publishing yield curves because of the great dispersion of yields at each maturity. There are obviously, therefore, grave limitations to the figures employed in this connection here.

2. This was done only once (November 1960) because of the difficulty of obtaining the relevant information. Since these series are maintained in such a way as to obtain a maximum degree of period-to-period comparability, there is no reason to believe that the results from this are unrepresentative.

bonds in the two series for the type of security in question. Thus, for example, since the component bonds of the McLeod, Young, Weir industrial bond yield series averaged fifteen years to maturity and those of the Moody's industrial "Aa" bond yield series averaged twenty-five years to maturity, the final Canadian figure is obtained by adding to the original McLeod, Young, Weir yield the yield on a twenty-five-year Government of Canada bond less that on a fifteen-year Government of Canada bond. The results are Canadian bond yield series which refer, roughly, to bonds of the same term as their corresponding U.S. yield series.[1] The public utility and industrial bond yield series refer to approximately 25 year bonds; the municipal and provincial (state) series refer to approximately 20 and 10 year bonds respectively.

The stock yield statistics are quite straightforward since they have been taken unaltered from their original sources. The Canadian figures (common and preferred) are compiled monthly by Moss, Lawson and Company. The U.S. common stock yield figures are taken, as are the bond yields, from publications of Moody's Investors' Service, while the preferred stock yield statistics are taken from those of Standard and Poor's.

1. Correlation Between U.S. and Canadian Interest Rates

Table 1, which shows simple correlation coefficients between U.S. and Canadian yields on equivalent bonds and stocks over the 1952–1960 period, summarizes the initial findings with respect to the first question.[2] Column 1 shows the degree of correlation between the interest rates, while column 2 shows that between first differences of the same interest rates. The latter was calculated in order to eliminate some of the effects of the autocorrelation which is present both in the U.S. and the Canadian series but, as it turned out, the use of first differences rather than absolute values had very little effect upon the value of any individual correlation coefficient.

From this table, one can see the remarkably high degree of correlation which existed between movements in U.S. and Canadian interest rates. Even the correlation between preferred stock dividend yields, which is far smaller than any of the others, is significantly different from zero at the

---

1. This procedure rests upon the assumption that the term structure of Government of Canada bond yields adequately represents that of yields on Canadian bonds issued by other borrowers. The decision to adjust the Canadian rather than the U.S. series in this respect was based upon the thesis, partially substantiated by the work of Durand and Winn [David Durand and Willis J. Winn, Basic Yields of Bonds 1926–47: Their Measurement and Pattern, NBER Technical Paper 5 (New York: National Bureau of Economic Research, 1947)], that differences in term structure are more likely among the sub-markets of the United States capital market, where the tax exemption factor plays an important divisive role, than among those of the Canadian market.

2. There were 107 observations, the first for January 1952 and the last for November 1960.

TABLE 1

Simple Correlation between U.S. and Canadian Interest Rates, 1952–1960

	Simple Correlation Coefficients	
	Absolute Values	First Differences
Federal Government 3 month Treasury bill	+.846	+.846
Federal Government 3 year bond	+.894	+.893
Federal Government 5 year bond	+.905	+.904
Federal Government 10 year bond	+.900	+.901
Provincial (State) Government 10 year bond [1]	+.862	+.865
Municipal Government 20 year bond [1]	+.860	+.866
Industrial 25 year bond [1]	+.952	+.954
Public Utility 25 year bond [1]	+.954	+.952
Preferred stock dividend yield	+.375	+.385
Common stock dividend yield	+.937	+.937

1. Term to maturity is approximate.

one per cent level of significance. One interesting implication of these results is that investors and borrowers, whether U.S. or Canadian, stood to gain very little in terms of diversification of risk from investing in or issuing foreign securities. Any such gain was almost certainly wiped out by the increased exchange risk thereby incurred.

The fact that all of these various types of security yields were highly correlated need not imply that they were in each case directly related to their foreign counterparts. In fact, domestic interest rates were themselves highly correlated with one another, so that it becomes difficult to ascertain precisely where the international ties, if any, originated.[1] The usual caution in the interpretation of correlation coefficients must also be exercised both in this and in subsequent tables. It may be, for example, that capital markets were affected by similar or related exogenous factors, such as the state of business, without any causal connections whatsoever existing between the capital markets themselves. Analysis of various types of long-term international

1. See Table 2.

capital movements indicates that direct connections between U.S. and Canadian capital markets did underlie these results.[1] It may be possible, nevertheless, to draw inferences from the interest rate data alone.

One could begin by comparing the correlation coefficient between, say, U.S. and Canadian industrial bond yields with those between the Canadian industrial bond yield and other Canadian bond yields, and between the U.S. industrial bond yield and other U.S. bond yields. From such a comparison, one might legitimately make inferences as to the relative importance of domestic and international influences upon any individual type of interest rate. The different terms to maturity of the bonds employed in the various series here being used complicate somewhat the interpretation of the correlation results but should not seriously influence them. Table 2, which makes comparisons of this sort, indicates, for instance, that federal, provincial (or state), and municipal government bond yields were noticeably more closely related to other domestic interest rates than they were to their foreign equivalents. Canadian public utility bond yields were more closely related to their U.S. counterparts than to domestic federal and provincial government bond yields, but their correlation with domestic municipal and industrial bond yields exceeded that with U.S. public utility bond yields. Substituting "industrial" for "public utility" and vice versa in the previous statement, it still holds. U.S. public utility bond yields' relationship with their Canadian equivalents was closer than that with any of the domestic U.S. yields except that on industrial bonds. U.S. industrial bond yields, on the other hand, were more closely related to domestic yields, with the single exception of those on state bonds, than to their Canadian counterparts.

This table shows clearly the high degree of correlation among domestic interest rates which renders it so difficult to disentangle international from domestic relationships. However, it provides little more than a hint as to the relative strengths of international and domestic interest rate ties. It appears that the strength of international ties was greater in the cases of industrial and public utility bond yields than in any of the others.

Table 3, which undertakes to show similar comparisons for stock dividend yields, suggests that preferred stock yields in the U.S. are closely related to industrial bond yields whereas this is not the case in Canada. On the other hand, in Canada, preferred stock yields are more closely related to common stock yields than they are in the U.S. (and this closer relationship, unlike that in the U.S., is a positive one). These figures suggest that the relatively poor correlation between U.S. and Canadian preferred stock yields observed in Table 1 results from an international difference in the characteristics of the preferred stocks used in the compilation of the two series. In Canada they appear to be substitutes for common stocks, whereas in the U.S. they are substitutes for bonds.

---

1. See below, Part B.

TABLE 2

Simple Correlations between U.S. and Canadian Bond Yields, 1952-1960 (Absolute Values)

	CANADA					UNITED STATES				
CANADA	Fed. Govt. 10 yr.[1]	Prov. Govt. 10 yr.[1]	Mun. Govt. 20 yr.[1]	Indus-trial 25 yr.[1]	Public Utility 25 yr.[1]	Fed. Govt. 10 yr.[1]	State Govt. 10 yr.[1]	Mun. Govt. 20 yr.[1]	Indus-trial 25 yr.[1]	Public Utility 25 yr.[1]
Fed. Govt. 10 yr.	-	+.983	+.916	+.916	+.916	+.900	-	-	-	-
Prov. Govt. 10 yr.[1]	+.983	-	+.938	+.926	+.925	-	+.862	-	+.895	+.904
Mun. Govt. 20 yr.[1]	+.916	+.938	-	+.969	+.977	-	-	+.860	+.905	+.928
Industrial 25 yr.[1]	+.916	+.926	+.969	-	+.993	-	-	-	+.952	-
Public Utility 25 yr.[1]	+.916	+.925	+.977	+.993	-	-	-	-	-	+.954
U.S.										
Fed. Govt. 10 yr.	+.900	-	-	-	-	-	+.915	+.931	+.962	+.936
State Govt. 10 yr.[1]	-	+.862	-	-	-	+.915	-	+.968	+.936	+.925
Mun. Govt. 20 yr.[1]	-	-	+.860	-	-	+.931	+.968	-	+.955	+.940
Industrial 25 yr.[1]	-	+.895	+.905	+.952	-	+.962	+.936	+.955	-	+.986
Public Utility 25 yr.[1]	-	+.904	+.928	-	+.954	+.936	+.925	+.940	+.986	-

1. Term to maturity is approximate.

TABLE 3

Simple Correlations between U.S. and Canadian Stock Dividend and Industrial Bond Yields
1952–1960
(Absolute Values)

	CANADA			UNITED STATES		
CANADA	Common stock	Preferred stock	Ind. 25 yr. bond[1]	Common stock	Preferred stock	Ind. 25 yr. bond[1]
Common stock	-	+.643	-.366	+.937	-	-
Preferred stock	+.643	-	+.319	-	+.375	-
Ind. 25 yr. bond[1]	-.366	+.319	-	-	-	+.952
U.S.						
Common stock	+.937	-	-	-	-.515	-.752
Preferred stock	-	+.375	-	-.515	-	+.943
Ind. 25 yr. bond[1]	-	-	+.952	-.732	+.943	-

1. Term to maturity is approximate.

While this approach has offered further insights into the question under consideration, perhaps a still more fruitful approach is one involving the use of partial correlation coefficients. The partial correlation coefficients shown in Table 4 vary in meaning. The coefficients are partial with respect only to the other yields grouped under the same letter (A, B, or C). Thus those in Section A are obtained by netting out the effects of the yields on other federal government obligations of different terms to maturity; those in Section B eliminate the effects of yields on obligations issued by other types of borrowers; and Section C's partials are with respect only to yields on stocks of both kinds and industrial bonds.

Most interesting in this table are the coefficients of partial correlation found in Section B. Since the market for federal government securities both in the U.S. and in Canada is the deepest and most active of all the submarkets in the capital market, one would expect that its international sensitivity would also be the greatest. Although the partial correlation between the yields on federal government 10 year bonds is significant at the one per cent level, it is far smaller than those between the yields on provincial (state) and public utility bonds. The high degree of partial correlation between Canadian provincial security and U.S. state security yields is particularly surprising. Recent work has shown that "There can be little doubt

TABLE 4

Partial Correlations between U.S. and Canadian Interest Rates, 1952–1960

(Absolute Values)

A.	
Federal Government 3 month Treasury bill	-.069
Federal Government 3 year bond	+.153
Federal Government 5 year bond	+.141
Federal Government 10 year bond	+.272
B.	
Federal Government 10 year bond	+.294
Provincial (State) 10 year bond[1]	+.609
Municipal Government 20 year bond[1]	-.151
Industrial 25 year bond[1]	+.205
Public Utility 25 year bond[1]	+.443
C.	
Preferred stock dividend yield	+.399
Common stock dividend yield	+.814
Industrial 25 year bond[1]	+.653

1. Term to maturity is approximate.

of the conclusion that the factor of tax exemption has isolated the market for state and local government securities from other capital markets."[1] One would expect that precisely the same tax considerations would produce a similar independence between these markets and their Canadian equivalents. Indeed, just such a lack of relationship is found in the case of municipal bond yields, and this makes the high correlation between provincial and state bond yields all the more puzzling. The partial correlation between industrial bond

1. Roland I. Robinson, "Factors Accounting for the Sharply Increased Cost of State and Local Government Borrowing," Journal of Finance, XII (May 1957), p. 130. See also, Robinson, Postwar Market for State and Local Government Securities (Princeton: National Bureau of Economic Research, 1960).

yields in the two countries, while not as high as the simple correlation, remains significant at the five per cent level.

Section A suggests that the strongest international ties between federal government bond yields come with longer terms to maturity; the partial correlation coefficient between 10 year federal government bond yields is significant at the one per cent level.

The dividend yield partial correlation coefficients in Section C have not had as many separate factors netted out. Their greater size relative to those in Sections A and B is therefore somewhat spurious. Nevertheless, both preferred and common stock dividend yields' international correlations remain highly significant in the statistical sense.

Supplementing Table 4 is Table 5, the analogue of Table 2, showing partial rather than simple correlations among various types of domestic bond yields. In the case of Canada, most striking are the high coefficients between federal and provincial government bond yields and between industrial and public utility bond yields. The fact that in both these cases the yields refer to securities of identical terms to maturity undoubtedly has boosted the values of the coefficients relative to the others in the table, but it cannot alone account for their height.[1] With those exceptions, Canadian federal government, provincial government, industrial, and public utility bond yields are all more closely related to their U.S. equivalents than to other domestic yields. In the U.S., relatively high domestic partial correlations are observed between industrial and public utility, state and municipal, and industrial and federal government bond yields. With these exceptions, federal government, state government, and public utility bond yields are more closely related to their Canadian equivalents than to other domestic U.S. yields.

Table 6 is analogous to Table 3, showing partial instead of simple correlation coefficients. From this table, it appears that Canadian preferred stock yields are not, after all, significantly more closely related to common stock yields than are their U.S. equivalents. In fact, the use of partials indicates that Canadian preferred stock yields are more closely related to U.S. preferred stock yields than to domestic common stock yields. It remains true that common stock and industrial bond yields are more

---

1. It could be objected that the yield figures might have offered more useful results had they been further adjusted so as to make all of them refer to the same term to maturity. This is a valid objection. No attempt at such adjustment was made because it was the primary object of the study to focus upon international rather than domestic equivalence of term, and because yield curves, at least in the U.S., vary among different classes of security and any resulting figures would therefore be of questionable reliability. See David Durand and Willis J. Winn, Basic Yields of Bonds 1926-47: Their Measurement and Pattern, N.B.E.R. Technical Paper 6 (New York: National Bureau of Economic Research, 1947).

TABLE 5

Partial Correlations between U.S. and Canadian Bond Yields, 1952–1960 (Absolute Values)

	CANADA					UNITED STATES				
CANADA	Fed. Govt. 10 yr.	Prov. Govt. 10 yr.[1]	Mun. Govt. 20 yr.[1]	Industrial 25 yr.[1]	Public Utility 25 yr.[1]	Fed. Govt. 10 yr.	State Govt. 10 yr.[1]	Mun. Govt. 20 yr.[1]	Industrial 25 yr.[1]	Public Utility 25 yr.[1]
Fed. Govt. 10 yr.	-	+.911	-.096	-.108	-.017	+.294	-	-	-	-
Prov. Govt. 10 yr.[1]	+.911	-	+.288	+.034	+.083	-	+.609	-	-.328	-.060
Mun. Govt. 20 yr.[1]	-.096	+.288	-	+.099	+.330	-	-	-.151	+.060	+.019
Industrial 25 yr.[1]	-.108	+.034	+.099	-	+.794	-	-	-	+.205	-
Public Utility 25 yr.[1]	-.017	+.083	+.330	+.794	-	-	-	-	-	+.443
U.S.										
Fed. Govt. 10 yr.	+.294	-	-	-	-	-	+.145	-.049	+.440	-.126
State Govt. 10 yr.[1]	-	+.609	-	-	-	+.145	-	+.691	+.163	+.327
Mun. Govt. 20 yr.[1]	-	-	-.151	-	-	-.049	+.691	-	+.122	-.206
Industrial 25 yr.[1]	-	-.328	+.060	+.205	-	+.440	+.163	+.112	-	+.638
Public Utility 25 yr.[1]	-	-.060	+.019	-	+.443	-.126	+.327	-.206	+.638	-

1. Term to maturity is approximate.

TABLE 6

Partial Correlations between U.S. and Canadian Stock Dividend and Industrial Bond Yields, 1952–1960

(Absolute Values)

	CANADA			UNITED STATES		
CANADA	Common stock	Preferred stock	Ind. 25 yr. bond[1]	Common stock	Preferred stock	Ind. 25 yr. bond[1]
Common stock	-	+.309	-.014	+.814	-	-
Preferred stock	+.309	-	+.438	-	+.399	-
Ind. 25 yr. bond[1]	-.014	+.438	-	-	-	+.653
U.S.						
Common stock	+.814	-	-	-	+.355	-.492
Preferred stock	-	+.399	-	+.355	-	+.825
Ind. 25 yr. bond[1]	-	-	+.653	-.492	+.825	-

1. Term to maturity is approximate.

closely related to their foreign equivalents than to other domestic yields in Section C. Furthermore, while U.S. preferred stock yields are still more closely related to yields on industrial bonds than are Canadian preferred stock yields, the international difference in this respect, though still substantial, is a little less than it appeared at first. Also of interest is the fact that the Canadian correlation coefficient between common stock and industrial bond yields, unlike that in the United States, is no longer significantly negative when partials are used.

2. Correlation between U.S. and Canadian Interest Rate Structures

In addition to correlation between the various interest rates themselves, correlation coefficients have also been calculated for differences between interest rates in order to shed some light upon the degree to which the structure of interest rates is related internationally. Table 7 uses the yields on federal government bonds of various terms to maturity to show those which relate to the term structure of rates.

High (simple) correlations are found to exist between Canadian and U.S. 10 year minus 3 year bond yields and between Canadian and U.S. 10 year minus 5-year yields. Those between U.S. and Canadian 10 year minus 3 month and 5 year minus 3 year yields, while lower, are still

TABLE 7

Simple Correlations between U.S. and Canadian Term Structure of Interest Rates, 1952–1960

Federal Government 3 year bond less Federal Government 3 month Treasury bill	-.165
" " 5 " " " " " 3 " " "	+.098
" " 10 " " " " " 3 " " "	+.327
" " 5 " " " " " 3 year bond	+.292
" " 10 " " " " " 3 " "	+.627
" " 10 " " " " " 5 " "	+.653

significant at the one per cent level. Correlations between 3 year minus 3 month yields and 5 year minus 3 month yields are not significant. Thus it appears that the further to the right one goes on the yield curve (that is, the higher the term to maturity), the greater is the correlation between the U.S. and the Canadian term structure.

The lower correlations on the left of the curve reflect the fact that yields on 3 month treasury bills (and other short-term instruments) are free to diverge to a greater extent from their foreign equivalents. This results from the existence of a forward exchange market,[1] and the greater values and variability of g_1, g_2, c, and x which are associated with the shorter time horizons implied by shorter terms to maturity.[2] These factors, incidentally, also explain the lower correlation coefficients observed with respect to 3 month treasury bills and shorter-term bonds in Tables 1 and 4. All in all, term structure does seem to be highly correlated in the segments of the yield curve where one would expect it to be so.

Table 8 presents similar correlation coefficients relating to the type-of-borrower structure of interest rates. All are significant at the one per cent level or lower. Again, however, most are far lower than the coefficients of correlation between the interest rates themselves. Extremely high coefficients were found for the relations between common stock, preferred stock, and industrial bond yields. As might have been expected on an a priori basis, the best bond yield correlations in this table are those between structural elements not involving provincial (state) or municipal securities.

V. Divergences in Average Levels of Equivalent U.S. and Canadian Interest Rates

The most important fact with respect to the relative levels of U.S. and Canadian interest rates during the 1952–1960 period is that all of the Canadian interest rates under consideration are higher than their U.S. equivalents. This is shown in Table 9, which presents the 9-year means of the monthly series introduced above. In the previous section it was seen that probably the most important factors in the explanation of such consistent differentials are persistent expectations of a fall in the value of the Canadian dollar or a relative rise in Canadian interest rates, the belief by U.S. investors that Canadian securities involve a greater risk of default, the existence of Canadian withholding taxation on interest income which would be tax-exempt in the U.S., the existence of exchange risk and transactions costs (exchange commissions) for U.S. investors, a higher marginal productivity of capital in Canada than in the U.S., and a relatively stricter monetary policy in Canada. The relative weight to be assigned to each of these possibilities is difficult to determine.

---

1. See above, p. 352, note 1.

2. See Section II above.

TABLE 8

Simple Correlations between U.S. and Canadian Type-of-Borrower Structure of Interest Rates,[1] 1952–1960

Provincial (State) Government 10 year bond less Federal Government 10 year bond	+.366
Municipal Government 20 year bond " " " 10 " "	+.351
Industrial 25 year bond " " " 10 " "	+.498
Public Utility 25 year bond " " " 10 " "	+.435
Municipal Government 20 year bond less Provincial (State) Government 10 year bond	+.344
Industrial 25 year bond less Public Utility 25 year bond	+.635
Common stock dividend yield less Industrial 25 year bond yield	+.977
Common " " " " Preferred stock dividend yield	+.871
Preferred stock dividend yield less Industrial 25 year bond yield	+.940

1. Terms to maturity are approximate.

The differences in the size of the Canadian–U.S. differential, which are shown for various types of securities in Table 9, must similarly be explained by the varying relative importances of each of these factors as well as by the existence of certain "structural" differences between U.S. and Canadian capital markets. Clearly, for example, the high differential in the case of provincial (state) and municipal bond yields is largely accounted for by the fact that interest income from these securities is tax-exempt in the U.S. but not in Canada. Other possible explanations for differences between various U.S.–Canadian differentials are that U.S. investors regard the default risk associated with Canadian industrial bonds as relatively greater than that associated with Canadian Government bonds, or that they regard the prospects for "growth" in Canadian common stocks as exceptionally high relative to U.S. stocks. Perhaps, too, Canada's higher marginal productivity of capital has a greater influence upon industrial bond yields than upon other interest rates. It is clear that conclusive findings with respect to the relative importances of the various factors producing divergent long-run levels of interest rates would require further statistical and institutional study.

TABLE 9

U.S. and Canadian Average Interest Rates, 1952–1960

	Canada	U.S.	Canada less U.S.
Federal Government 3 month Treasury bill	2.539	2.279	+.260
Federal Government 3 year bond	3.580	2.963	+.617
Federal Government 5 year bond	3.848	3.095	+.753
Federal Government 10 year bond	3.869	3.199	+.670
Provincial (State) Government 10 year bond[1]	4.426	2.457	+1.969
Municipal Government 20 year bond[1]	5.000	3.144	+1.856
Industrial 25 year bond[1]	4.962	3.574	+1.388
Public Utility 25 year bond[1]	4.830	3.861	+.969
Preferred stock dividend yield	4.993	4.350	+.643
Common stock dividend yield	4.566	4.364	+.202

1. Term to maturity is approximate.

B. Portfolio Capital Movements as a Source of Connections between United States' and Canadian Capital Markets, 1952–1960

Canadian long-term interest rates have obviously been highly correlated with equivalent U.S. interest rates. Even interest rate structures, of both the term to maturity and type of borrower variety, showed significant international relationships. These high correlations could have resulted from remarkably similar responses on the part of the Federal Reserve System and the Bank of Canada to remarkably similar economic conditions during the 1952–1960 period. An explanation which is both more plausible and more appealing to the theorist, however, is that U.S. and Canadian capital markets were bound to one another by the unifying influences discussed in Section I of Part A. The most important such influence was seen to be the international flow of long-term securities. Through investigation of the roles which the more volatile of the elements affecting international investment and borrowing decisions played in the actual movement of portfolio capital between the U.S. and Canada during the period, it is possible to test this theoretical explanation of the high correlations between interest rates.

For the purpose of discussion, long-term capital movements may be broken down into sales and purchases to and from foreigners of new issues of securities and outstanding securities. Each will be discussed in turn. (Retirement of foreign-held securities are another form of long-term capital movement but these are of less theoretical interest since they arise largely through the mere passage of sufficient lengths of time. In some instances below, however, it will be convenient to deal with net new security issues, i.e., gross new issues less retirements.)

In order to place the subsequent discussion in its proper perspective, it may be useful to present a broad summary picture of long-term portfolio capital movements between the U.S. and Canada during the 1952–1960 period. Such a picture may be obtained from Table 10. It will at once be noticed that there are certain elements of recent experience which present a reasonably consistent pattern. Trade in outstanding Canadian bonds and debentures has resulted in capital outflows from Canada to the United States in every year but one; that is to say, Canadians have generally purchased more outstanding Canadian bonds from U.S. residents than they have sold to them. On the other hand, outstanding Canadian stocks have been sold by Canadians in greater values than they have been purchased in trade with U.S. residents. The same is true of outstanding United States stocks. Net new issues of Canadian securities (proceeds of new issues minus retirements) have produced very substantial capital inflows to Canada in every year but one, whereas Canadian residents' purchases of U.S. new issues (net) have consistently resulted in small capital outflows from Canada. In sum, long-term portfolio capital transactions between Canada and the United States have resulted in capital inflows to Canada in every year of the period except 1955.

TABLE 10

Portfolio Capital Transactions between Canada and the United States, 1952–1960[1]

(Millions of Dollars)

	1952	1953	1954	1955	1956	1957	1958	1959	1960
Canadian Securities:									
Trade in outstanding bonds and debentures	-163	- 85	- 86	-164	- 36	- 60	- 6	+ 24	- 7
Trade in outstanding common and preferred stocks	+ 65	+ 6	+ 85	+ 95	+ 76	+ 9	+ 68	+ 76	+ 54
Proceeds of new issues	+315	+303	+299	+127	+601	+722	+600	+623	+332
Retirements	- 75	-134	-184	-211	-134	-105	-132	-211	-173
Total	+142	+ 90	+114	-153	+507	+566	+530	+512	+206
U.S. Securities:									
Trade in outstanding bonds and debentures	-	+ 9	+ 6	+ 1	+ 3	- 12	+ 10	- 15	+ 11
Trade in outstanding common and preferred stocks	+ 7	+ 10	+ 8	+ 28	+ 22	+ 44	+ 30	- 2	+ 8
Proceeds of new issues	- 5	- 18	- 3	- 8	- 13	- 15	- 13	- 8	- 15
Retirements	-	+ 1	+ 1	+ 2	-	+ 4	+ 4	+ 6	+ 12
Total	+ 2	+ 2	+ 12	+ 23	+ 12	+ 21	+ 31	- 19	+ 16
Total	+144	+ 92	+126	-130	+519	+587	+561	+493	+222

1. Signs indicate the direction of capital flows: (+) for inflows into Canada, (-) for capital outflows from Canada.

Source: The Dominion Bureau of Statistics.

I. The Independent Variables

1. Interest Rate Differentials

Portfolio capital can be expected to flow at varying rates across the international border in response to varying differentials between equivalent Canadian and U.S. interest rates. The most obvious independent variables to be employed are therefore international interest rate differentials of various types. These differentials will always be expressed as the Canadian interest rate (R_C) less the equivalent U.S. interest rate (R_U). A description of the interest rate data to be used for this purpose has been presented above. These raw data refer to month-ends. When the regressions in question employ monthly capital flow data, the monthly interest rate differentials used are the averages of those in two successive month-ends. When the dependent variable (capital flow) data are quarterly, the averages of interest rate differentials in four successive month-ends are employed as quarterly data.

2. Exchange Rate Expectations

As has been shown above, another potentially important determinant of international portfolio capital flows is the state of expectations as to the exchange rate (g_1). This extremely volatile element can be expected to play a greater role the shorter is the time horizon of the seller or purchaser in question, since high annual rates of return may be earned through relatively small shifts in the exchange rate.

The measurement of expectational variables is an impossibility. The best one can hope for is the obtaining of a reasonable proxy. Where arbitrage between money markets and the forward market for foreign exchange is imperfect, as in the U.S.-Canadian case, it is possible to employ the forward discount or premium on one currency, adjusted for the international differential in interest rates of the same (short) term, in this way. Changes in the value of such a variable can be expected to reflect changes in the market's estimate of the need to cover against the risk of short-term fluctuations in the exchange rate when that particular currency is held. Individual market participants will, of course, react differently to changes in the value of this and any other proxy variables. Rhomberg employed, with some success, the period-to-period change in the rate of exchange as another proxy for short-term exchange expectations in his attempts to explain the speculative demand for Canadian dollars.[1] Attempts to incorporate this variable into the regression analysis below, however, yielded consistently insignificant results. A third possibility is to use the spot exchange rate itself as a proxy for more long-run expectations. Over the 1952–1960 period, such a procedure would have been justifiable, for the values of the U.S. and Canadian dollars relative to one another showed no particular trend. Thus,

1. Rudolf R. Rhomberg, "Canada's Foreign Exchange Market: A Quarterly Model," International Monetary Fund Staff Papers, VII (April 1960), pp. 442–43, 445, 450.

when the Canadian dollar's value was abnormally high, it might reasonably have been expected to fall; conversely, when its value was abnormally low, it could be expected to rise. The first and, in one instance, the third of these proxies are those actually employed in subsequent sections. To distinguish the two, they will be referred to as g_S and g_L (for short-term and long-term) respectively.

3. Divergent Interest Rate Expectations

Relative expectations as to the chances for purely domestic capital gains (that is, relative interest rate expectations) may also greatly influence international transactions in long-term securities, particularly where the transactors have short-term intentions. Culbertson has calculated that annual rates of return of 23% and 97% could have been earned through the holding of U.S. Government bonds for three-month and one-week periods, respectively, during 1953.[1] For short-term transactors, small divergences in expectations for U.S. and Canadian interest rates may therefore offer substantial inducements for international transactions in long-term bonds. How to measure interest rate expectations presents a problem. It is dealt with here by the assumption that the observed Canadian–U.S. interest rate differential ($R_c - R_u$) is a valid measure not only of relative long-term yields, but also of relative yields in the shorter-run allowing for the possibility of capital gains or losses on long-term securities. This assumption rests on the premise that there exists a "normal" Canadian–U.S. interest rate differential for each class of security. From this it follows that when the Canadian–U.S. interest rate differential is abnormally large, it can be expected ultimately to fall. In order for this to occur, the Canadian rate must rise less or fall more than its U.S. counterpart, thus in either case giving rise to the prospect of smaller capital losses or greater capital gains in Canadian securities than in their U.S. equivalents. It will therefore pay an investor to move his funds into Canadian (U.S.) securities when the Canadian interest rate relative to the U.S. equivalent is abnormally high (low), regardless of whether he is an income-conscious or a capital-conscious investor. Unfortunately, it is not possible to separate the income motives from the capital gains motives through the use of one independent variable which reflects both.

II. New Issues

Table 10 provides an indication of the very great importance of foreign new issues relative to the other components of the total long-term capital flow between the two countries. There are four major classes of inter-capital market transactions which are of interest in connection with these new security issues: 1) sales of Canadian new issues payable in U.S. dollars to U.S. investors, 2) sales of Canadian new issues payable in

1. J. M. Culbertson, "The Term Structure of Interest Rates," Quarterly Journal of Economics, LXXI (November 1957), pp. 506–07.

Canadian dollars to U.S. investors, 3) sales of U.S. new issues payable in U.S. dollars to Canadian investors, and 4) sales of U.S. new issues payable in Canadian dollars to Canadian investors.[1] Since the U.S. capital market is much broader, deeper, and more developed than the Canadian, and interest rates on it have consistently fallen short of their Canadian counterparts, instances of the fourth type have been rare.[2] This study will therefore confine itself to the other, more common, classes of new issue capital movements. During the 1952–1960 period it was the first of these three types, i.e., sales of Canadian new issues payable in U.S. dollars to U.S. investors, which was by far the most important in terms of size and volatility. Next in importance were sales of Canadian new issues of both bonds and stocks payable in Canadian dollars to U.S. investors. Capital flows in the reverse direction in the form of Canadian purchases of U.S. issues (payable in U.S. dollars) were relatively small. Each of these types of international new issues will be considered separately below.

The most comprehensive figures on international new issues are those collected by Canada's Dominion Bureau of Statistics.[3] These statistics, since they are collected in connection with the preparation of balance of payments estimates, all apply to the period during which new issues, retirements, and sales and purchases of outstanding issues are delivered, that is, the period during which the currency actually changes hands. From the point of view of this study's attempt to relate flows to other independent factors, this period is, unfortunately, not the relevant one. Rather, what would be most useful are data which refer to the period during which the decision to transact was made. This limitation is not of great importance as far as long-run over-all views of the composition of capital flows are concerned. In any case, since these data are the best available they have been, and will be, used for this purpose. Nor is it even of particular importance where more detailed analysis of the trade in outstanding securities is concerned, for the length of time elapsing between contract and delivery

---

1. It is conceivable that Canadian new issues payable in U.S. dollars are at times sold to Canadian investors (or U.S. issues payable in Canadian dollars are sold to U.S. investors). In such cases, both the borrower and the investor have entered the foreign capital market since both assume exchange risk. The foreign capital market has then served as an intermediary between domestic borrowers and investors.

2. At least one such instance did occur during the period. Dominion Bureau of Statistics, The Canadian Balance of International Payments (1955), p. 24.

3. Dominion Bureau of Statistics, The Canadian Balance of International Payments (quarterly and annually, 1952–1960), and Purchases and Sales of Outstanding Securities between Canada and Other Countries (monthly, 1952–1960).

dates in this trade rarely exceeds three or four days, whereas the data periods under discussion are a month long. However, as recent D.B.S. work has demonstrated, the use of quarterly delivery date statistics for new security issues may produce results which are substantially different from those which would be obtained through the use of contract date figures. In the years 1958–1959, during which quarterly values for total Canadian new issues sold to the U.S. varied between 15 and 200 million dollars, it was estimated that the divergence between contracted and delivered values for this variable varied between zero and 73 million dollars.[1] Thus the validity of the analysis would be greatly augmented by the use of the relevant (contract date) data. This has been possible for only one of the three types of international new issues, fortunately the most important one. Unpublished material compiled by A. E. Ames and Company was available, permitting the preparation of quarterly contract date statistics for new issues by Canadians payable in U.S. dollars. In that these statistics exclude refunding operations and are stated at par in terms of U.S. dollars, their meaning is slightly different from that of the D.B.S. estimates.[2] In the relevant regression analysis below, both the D.B.S. delivery date data and these Ames contract date data have been employed.

An obstacle to the employment of conventional statistical analysis is the "lumpiness" of international new issues in the Canadian case. Since the market is a thin one, quarter-to-quarter changes may often be extremely large and one large issue may easily outweigh all others during the period. This obstacle, which will result in low values for the coefficient of determination (R^2) and, possibly, less significant slope coefficients in the calculated regression equations, exists whether one is concerned with Canadian issues payable in U.S. dollars, Canadian dollar issues sold to U.S. investors, or U.S. dollar issues sold to Canadians. An attempt was made, without success, to get around this difficulty by employing the number of separate new issues by Canadians in the U.S. market, as obtained from A. E. Ames and Company, rather than the value of these issues as the dependent variable in the relevant regression equations.

1. Canadian New Issues Payable in U.S. Dollars

The incentive for the flotation of Canadian new issues payable in U.S. dollars clearly comes from the side of the issuer of the securities, since it is he who has to enter the foreign capital market. All levels of Canadian Government as well as Canadian corporations have at one time or another sold new issues of bonds or debentures which are payable in U.S. dollars to

1. Dominion Bureau of Statistics, The Canadian Balance of International Payments, 1959, p. 37.

2. D.B.S. estimates are stated as the Canadian dollar receipts of the issuers, net of discounts, underwriting commissions, and other expenses. Dominion Bureau of Statistics, The Canadian Balance of International Payments, 1959, p. 38.

U.S. investors; this may be seen in Table 11. During the 1952–1960 period, however, the Government of Canada did not issue any securities of this type. This lack of Federal Government response to incentives which induced lower levels of government and corporations to enter the U.S. capital market for substantial funds undoubtedly reflects the views of the Bank of Canada's administration. "For provinces and municipalities to borrow abroad is hazardous to themselves and undesirable in the national interest," in the view of the then Governor of the Bank.[1] This outlook has not always governed the management of Federal public finance in Canada, with the result that there still exist outstanding Government of Canada bonds payable in U.S. dollars; their retirement from time to time during the period has produced the negative net new issues contained in Table 11. Since the decisions of the Bank of Canada appear to be based upon different criteria from those employed by Canadian borrowers, the totals in this and subsequent tables exclude Government of Canada issues.

What is of theoretical interest is not the absolute amount issued in the U.S. market, as shown in Table 11, for this may be affected by many exogenous factors which also affect the amount of domestic security issues, but the amount issued in the U.S. market relative to the amount issued in the Canadian market. The regression analysis to follow will therefore focus upon the relative degree of Canadians' resort to U.S. capital markets, as measured by the percentage of total new issues (gross), payable either in Canadian or U.S. dollars, which are payable (solely or optionally) in U.S. dollars.

The relative extent of resort to this type of financing was related not only to interest rate differentials and exchange rate expectations, but also to the size of the issues to be floated. The relative role of the U.S. capital market in the total of new issues of any particular size increases with the size of the issues. Although issues of amounts totaling less than 600 thousand dollars (an arbitrary cutoff point) accounted for a very large share of the total number of separate trips to the market during the 1952–1960 period, the place of U.S. dollar securities in these amounts was negligible. The contract date (Ames) data was therefore adjusted so as to include only issues of more than this amount. Detailed information permitting a similar adjustment in the delivery date (D.B.S.) data was unfortunately not available.

Quarterly data for 1952–1960[2] on the percentage of total U.S. and Canadian dollar issues by Canadian borrowers accounted for by those

---

1. Annual Report of the Governor of the Bank of Canada to the Minister of Finance, 1960, p. 19.

2. For technical reasons, the delivery date (D.B.S.) observations began in the second quarter, whereas the contract date (Ames) data began in the first quarter of 1952. Both continued to the fourth quarter of 1960.

TABLE 11

New Issues (Gross and Net) of Canadian Bonds and Debentures Payable (Solely or Optionally) in U.S. Dollars, by Type of Borrower, 1952–1960

(Millions of Dollars)

	1952	1953	1954	1955	1956	1957	1958	1959	1960
New Issues (Gross):									
Government of Canada[1]	-	-	-	-	-	-	-	-	-
Provincial Governments[1]	101	136	114	-	212	130	162	316	84
Municipal Governments	43	69	34	40	106	116	143	147	110
Corporations	117	29	96	5	169	388	185	49	70
Total	260	234	244	45	487	634	490	512	265
New Issues (Net):									
Government of Canada[1]	-	-74	-21	-94	-62	-19	-	-17	-
Provincial Governments[1]	66	107	57	-19	198	107	127	281	42
Municipal Governments	28	56	11	22	89	97	117	120	80
Corporations	106	19	67	-53	153	370	161	7	14
Total excluding Govt. of Canada	199	181	135	-50	439	574	405	408	137

1. Direct and guaranteed.

Source: The Dominion Bureau of Statistics.

payable in U.S. dollars (both in value terms) have been employed as dependent variables in linear regression equations in which the relevant interest rate differentials ($R_c - R_u$) and the earlier-mentioned measure of short-term exchange rate expectations (g_S) were employed as independent variables. The other proxy for exchange rate expectations (g_L, the spot rate) is suitable for use only in connection with the trade in outstanding securities, because it is itself affected by changes in the volume of foreign new issues, which bulk large in the absolute total of international capital flows. Both the delivery date and the contract date data were tried for each type of borrower's new issues.[1] More than one interest rate differential was tested for each type of borrower's issues (but not for total issues).

Since the interest rate differential variable was stated as the Canadian yield minus the U.S. yield ($R_c - R_u$), the desired sign on this variable is positive. The short-term exchange rate expectation variable (g_S) was stated as the percentage three-month forward premium on the U.S. dollar less the difference between the Canadian and the U.S. three-month Treasury bill rates (all expressed as yields per annum). Thus, the higher its algebraic value, the greater is the market's expectation that the U.S. dollar will rise in value relative to the Canadian dollar. It follows that the desired sign associated with this variable is negative.

Table 12 shows the results of these regression calculations and serves as the model for all subsequent presentations of regression results. Only those results are shown in which at least one of the slope coefficients of an equation differs from zero at the ten per cent level of significance or lower, using a one-tail test of the null hypothesis. In the left-hand column, the interest rate differential used in the equation in question is shown. Where the slope coefficients of both of the independent variables are significantly different from zero (as shown by the asterisks), the beta coefficients have been computed.

---

1. To obtain delivery date (D.B.S.) figures for total United States and Canadian dollar issues by various types of Canadian borrowers, a little manipulation of the published data was required. From total deliveries in all currencies, it was necessary to subtract all those payable in foreign currencies, including some payable in pounds sterling, Swiss francs, etc., and then add back those payable in U.S. dollars.

The figures for 1952 to 1954 required still a little more work. Total municipal issues were available only on an annual basis during this period. The Canadian dollar component for the year, arrived at as described above, was therefore prorated over the year's four quarters and the quarterly data on U.S. dollar issues was then added to these amounts. A recent revision in the annual totals for provincial and corporate new issues, without the publication of the revised quarterly data, also caused some trouble in this period's figures. The differences between the original annual totals and the revised annual totals for the corresponding years were prorated over the four quarters.

TABLE 12

Regression Results—Dependent Variables: Percentage of Total U.S. and Canadian Dollar New Issues by Canadians Payable (solely or optionally) in U.S. Dollars, by Type of Borrower, Quarterly 1952–1960

	Constant	$R_c - R_u$	g_S	R^2
		$H_o: b_1 = 0$	$H_o: b_2 = 0$	
		$H_1: b_1 = (+)$	$H_1: b_2 = (-)$	
Provincial Governments				
Contract Dates (Ames):				
Federal Govt. 10 year	+ 9.376	+ 29.614**	- 8.945*	.096
	(10.416)[1]	(16.509)	(6.404)	
		B = 3.45	B = 2.69	
Delivery Dates (D.B.S.):				
Federal Govt. 10 year	+ 3.563	+ 31.073***	- 6.200*	.188
	(7.170)	(11.406)	(4.405)	
		B = 5.02	B = 2.59	
Prov.(State) Govt. 10 year[2]	- 2.751	+ 12.291*	- 1.609	.073
	(15.150)	(7.778)	(4.172)	
Industrial 25 year[3]	- 3.430	+ 39.435***	- 1.410	.181
	(2.084)	(14.795)	(3.852)	
Public Utility 25 year[3]	- 3.499	+ 25.903*	- 2.117	.072
	(15.443)	(16.209)	(4.247)	
Municipal Governments				
Contract Dates (Ames):				
Federal Govt. 10 year	+ 8.883	+ 30.963**	- 0.757	.116
	(10.554)	(16.729)	(6.489)	
Municipal Govt. 20 year[3]	- 7.296	+ 17.797*	+ 5.220	.092
	(21.422)	(11.332)	(5.657)	
Public Utility 25 year[3]	- 1.500	+ 29.601*	+ 3.297	.073
	(21.422)	(22.527)	(5.930)	
Delivery Dates (D.B.S.):				
Municipal Govt. 20 year[3]	- 8.759	+ 15.619**	+ 1.735	.099
	(16.150)	(8.512)	(4.062)	

Table 12 (continued)

	Constant	$R_c - R_u$	g_S	R^2
Corporations				
Contract Dates (Ames):				
Federal Govt. 10 year	- 0.692 (7.468)	+ 26.419** (11.837)	+ .165 (4.592)	.171
Delivery Dates (D.B.S.)				
Federal Govt. 10 year	+ 5.488 (4.858)	+ 12.549* (7.728)	+ 0.614 (2.984)	.111
Public Utility 25 year[3]	- 1.724 (5.960)	+ 9.517* (6.256)	- 0.900 (1.639)	.067
Total (excluding Govt. of Canada)				
Contract Dates (Ames):				
Federal Govt. 10 year	+ 3.777 (6.112)	+ 32.295*** (9.689)	- 4.262 (3.758)	.259
Delivery Dates (D.B.S.):				
Federal Govt. 10 year	+ 6.837 (4.727)	+ 19.814*** (7.521)	- 2.312 (2.904)	.186

*** Significantly different from zero at the 1 per cent level (one-tail test).
** " " " " " " 5 " " " " " "
* " " " " " " 10 " " " " " "

1. Bracketed values in this and subsequent tables indicate the relevant standard errors.

2. Direct and guaranteed.

3. Term to maturity is approximate.

It will be observed that all types of borrowers' new issues are found to respond to the Canadian–U.S. interest rate differential. This response is observable not only when the contract date (Ames) data are used, but also when the delivery date (D.B.S.) data are used as dependent variables. Indeed, oddly enough, the delivery date statistics at times yield more significant results than those relating to contract dates. Thus, the lag between contract and delivery dates does not seem greatly to impede analysis which is based upon delivery data alone. Of particular interest is the success attained through the use of the federal government yield differential: only in the case of municipal

deliveries of new issues does it fail to yield a significant slope coefficient. Furthermore, in all but the latter case, the statistical significance of the yield differential coefficient is greater when the federal government yields are employed than when any of the alternatives are used. This suggests that Canadian borrowers considered the general state of the capital markets, for which the federal government yield is probably the best proxy, rather than the state of particular sub-markets thereof when reaching their decision as to which capital market in which to float their issues. It may also be that the observed yields on sub-markets of the Canadian capital market do not accurately reflect the information about these markets which would be relevant to issuers' decisions; in particular, these sub-markets are all so thin that consideration must be given to the possible effects of a large new issue and the ease with which such an issue could be sold at the prevailing yields.[1]

The slope coefficient associated with short-term exchange rate expectations, g_S, is significant only in the provincial government equations, and even in these, its beta coefficient is well below that of the interest rate differential. Since this factor is not of great importance to the long-term borrower,[2] this lack of significant results is not surprising. The possibility remains, of course, that the proxy for this factor here being used is not a good one, and that exchange rate expectations are, in fact, far more important than it would appear from these results.

2. Canadian New Issues Payable in Canadian Dollars Sold to the U.S.

Unlike the U.S. dollar new issues of Canadian borrowers, the Canadian dollar issues sold to U.S. investors are best thought of as investor-motivated capital flows. In this case, it is the U.S. investor who steps outside of his domestic capital market. This interpretation may at times be misleading; for example, if a U.S. investor's purchase of Canadian dollar securities were the result of a U.S. sales campaign by the Canadian issuer, it might reasonably be said that at least part of the incentive for the international flow of capital arose from the Canadian side. In general, however, it seems the most useful one.

Table 13 shows the participation by U.S. investors in the Canadian new issues market. The amounts involved are far smaller than those arising from Canadian borrowers' resort to the U.S. capital market. On balance, during the 1952–1960 period, the U.S. has purchased (net) Canadian dollar obligations of provincial governments, municipal governments, and corporations; corporate bonds and stocks have accounted for the largest values purchased. At the same time, it has had retired larger amounts of Government of Canada securities than it has purchased in the Canadian new issues market.

Ideally, the analysis should proceed to an explanation of the changes in U.S. holdings of Canadian dollar securities, or even this change relative

1. See above, p. 357.

2. See above, pp. 354–55.

TABLE 13

New Issues (Gross and Net) of Canadian Securities Payable in Canadian Dollars to the U.S., by Type of Borrower, 1952–1960

(Millions of Dollars)

	1952	1953	1954	1955	1956	1957	1958	1959	1960
New Issues (Gross):									
Government of Canada[1]	-	-	1	-	7	11	45	27	13
Provincial Governments[1]	6	5	6	1	9	1	2	10	1
Municipal Governments	4	4	5	2	6	6	3	8	4
Corporations (bonds)	22	19	34	34	49	46	42	48	38
Total bonds and debentures	32	28	47	38	70	64	93	92	57
Corporations (stocks)	23	41	8	45	44	24	17	19	10
New Issues (Net):									
Government of Canada[1]	-11	- 1	-38	-	-10	10	38	-33	-10
Provincial Governments[1]	5	5	4	1	9	1	- 5	8	-
Municipal Governments	4	3	2	2	5	3	1	4	1
Corporations (bonds)	21	18	29	16	46	39	27	36	28
Total bonds and debentures	19	26	- 1	19	50	53	71	14	20
Corporations (stocks)	23	37	3	40	40	10	3	6	3

1. Direct and guaranteed.

Source: The Dominion Bureau of Statistics.

to total U.S. holdings of all comparable assets. This, however, would require data on the trade in outstanding securities, both with Canada and with other countries, classified according to currency of repayment, and these data are not available. The regressions which were run used net new Canadian dollar issues sold to U.S. investors as the dependent variables. To analyze this type of capital flow in such a "partial" manner is to ignore the undoubtedly quite important U.S. portfolio changes which are arranged through the trade in outstanding securities. Unfortunately, there was no availabe alternative procedure.

Therefore, as might have been expected, the regression results, tabulated in Table 14, are not as good as the previous ones. No significant coefficients could be found for net new Canadian dollar issues of the Government of Canada, total Canadian bonds and debentures, or total Canadian (common and preferred) stocks. The lack of response in the case of Canadian Government bonds probably reflects the fact that, up until the last four years or so, net new issues of this type were primarily accounted for by the passive response on the part of U.S. investors to the retirement of their holdings. The poor results obtained for U.S. investment in Canadian stocks is the product of the facts that much of this investment was induced by long-run growth prospects, which are not really measured by stock dividend yields, and that still more represented the building up of assets by a number of U.S. investment companies which were formed during this period for special tax treatment reasons.[1]

Nevertheless, other types of investment did result in significant slope coefficients for the interest rate differential variable. U.S. investors in Canadian dollar issues of provinces and municipalities apparently responded to provincial-state and federal government yield differentials, respectively, and their net purchases of new corporate bond issues were significantly related to both industrial and public utility bond yield differentials. In none of these equations was the coefficient of the exchange rate expectations variable found to be significantly different from zero. Again, if investors had long-term intentions, one would not expect too much in the way of response to exchange rate expectations.

3. Canadian Purchases of U.S. New Issues

The small amounts which Canadians invest in the U.S. new issues markets are shown in Table 15. A regression of the type described in the previous section was also run on this form of capital flow, using total net

1. J. Grant Glassco, Royal Commission on Canada's Economic Prospects, Certain Aspects of Taxation Relating to Investment in Canada by Non-residents (Ottawa: Queen's Printer, 1956), p. 19; Lawrence M. Greene, "The Investment Company—Useful Medium for American Investment Abroad," Journal of Finance, XIV (December 1959), p. 496; Dominion Bureau of Statistics, The Canadian Balance of International Payments, 1959, p. 36.

TABLE 14

Regression Results–Dependent Variables: Net New Issues of Canadian Dollar Securities[1] Sold to U.S. Investors, by Type of Borrower, Quarterly 1952–1960

	Constant	$R_c - R_u$	g_S	R^2
		$H_o: b_1 = 0$	$H_o: b_2 = 0$	
		$H_1: b_1 = (+)$	$H_1: b_2 = (-)$	
Government of Canada				
No significant results				
Provincial Governments[2]				
Prov.-State Govt. 10 year[3]	- 1.404	+ .967*	+ .494	.123
	(1.385)	(.711)	(.381)	
Municipal Governments[2]				
Federal Govt. 10 year	+ 2.839	+ .748*	- .190	.069
	(3.087)	(.491)	(.90)	
Corporations (Bonds)				
Industrial 25 year[3]	- 3.878	+ 4.111**	- .299	.131
	(2.660)	(1.889)	(.492)	
Public Utility 25 year[3]	- 1.244	+ 3.326**	- .420	.082
	(1.905)	(1.999)	(.524)	
Total Bonds and Debentures				
No significant results				
Corporations (Stocks)				
No significant results				

** Significantly different from zero at the 5 per cent level (one-tailed test).
* Significantly different from zero at the 10 per cent level (one-tailed test).

1. In millions of dollars.
2. Direct and guaranteed.
3. Term to maturity is approximate.

new issues purchased by Canadians in the U.S. as the dependent variable, but no significant slope coefficients were uncovered. This inconclusive result may be the product of the fact that, as in the previous section, the dependent variable should allow for the international trade in outstanding securities—in this case, trade in U.S. securities. Another important possibility is that the dependent variable figures, being very small, suffer from excessive rounding. On the other hand, it may be that Canadian investments in U.S. securities are governed by special factors, such as the desire for risk diversification, and would not show significant response to the independent variables being employed even with the best of dependent variable statistics.

TABLE 15

Canadian Purchases (Gross and Net) of U.S. New Security Issues, 1952–1960[1]

(Millions of Dollars)

	1952	1953	1954	1955	1956	1957	1958	1959	1960
New Issues (Gross)	-5	-18	-3	-8	-13	-15	-13	-8	-15
New Issues (Net)	-5	-17	-2	-6	-13	-11	- 9	-2	- 3

1. Signs indicate the direction of capital flows: (+) for inflows into Canada, (-) for capital outflows from Canada.

Source: The Dominion Bureau of Statistics.

III. The Trade in Outstanding Securities

In addition to the capital market ties deriving from the flotation of new Canadian issues in the United States and the new issue purchases of investors of both countries in foreign markets, there are those associated with the extensive international trade in outstanding issues of both bonds and stocks. The degree of international activity in the markets for outstanding securities, and therefore, presumably, their influence in tying together equivalent U.S. and Canadian yields, varies greatly, depending upon the type of security in question. By far the greatest volume of international trade is found in the markets for Canadian and U.S. stocks and Government of Canada securities. It is interesting that these particular types of securities were found in the previous section not to figure importantly in the achievement of international ties through the medium of the new issues markets. Next in volume of trade come corporate, provincial, and municipal (in that order) bonds and debentures.

The annual net trade in outstanding securities of all the various types

is shown in Table 16. As was seen earlier, trade in stocks resulted in a capital flow of the reverse direction from that in bonds and debentures. The Canadian net purchases of outstanding Canadian bonds and debentures have been spread through all of their types, except for the fact that in recent years sales of Government of Canada securities to the U.S. have exceeded purchases. The net trade in U.S. bonds and debentures has not exhibited as steady a pattern as that in Canadian securities. Whatever "steadiness" of behavior is evident in Table 16 is a little misleading, however, for day-to-day and month-to-month movement of this type of international capital flow is actually extremely irregular both in amount and in direction.

The factors underlying the extent and direction of this flow are somewhat more diverse and complex than those relating to the flow of new issue capital. There seems to be substantial agreement that international trade in outstanding securities is best looked upon as a short-term capital flow.[1] That this is not the entire story has also, however, been recognized. Much of this capital flow undoubtedly involves the purchase of long-term obligations by income-conscious, rather than capital-conscious, investors who intend to hold them until maturity or for long periods of time.[2] For this type of investor, the market in outstanding securities has the advantage of offering him a much wider selection of assets among which to choose than he could ever find in the new issue market at any particular moment. Still another component of the trade in outstanding securities may be transactions involving the original issuers of securities, which might better be included under "retirements."

As has already been mentioned, the fact that D.B.S. monthly data on the trade in outstanding securities concern the date of delivery rather than the date of the contract is not a serious limitation, since the lag between the two is not likely to be more than a few days. Nor are they subject to the "lumpiness" problems associated with the new issues data, for this trade occurs reasonably steadily from day to day. D.B.S. figures are

---

1. Arthur I. Bloomfield, "The Significance of Outstanding Securities in the International Movement of Capital," Canadian Journal of Economics and Political Science (November 1940), p. 517; R.A. Radford, "Canada's Capital Inflow, 1946-53," International Monetary Fund Staff Papers, IV (February 1955), pp. 226-27; Donald Bailey Marsh, World Trade and Investment (New York: Harcourt, Brace and Co., 1951), p. 59; Richard E. Caves and Richard H. Holton, The Canadian Economy, Prospect and Retrospect (Cambridge: Harvard University Press, 1959), pp. 109-10; Paul Wonnacott, The Canadian Dollar, 1948-1958 (Toronto: University of Toronto Press, 1960), p. 124.

2. R. A. Radford, op. cit., pp. 226-27, 238; Dominion Bureau of Statistics, The Canadian Balance of International Payments, A Study of Methods and Results (Ottawa, 1939), pp. 108, 132-33.

TABLE 16

Canadian Sales Minus Purchases of Outstanding Securities to the U.S., by Types, 1952-1960

(Millions of Dollars)

Bonds and Debentures	1952	1953	1954	1955	1956	1957	1958	1959	1960
Canadian:									
Government of Canada[1]	-157	- 64	- 69	- 84	- 21	- 16	20	45	20
Provincial Governments[1]	- 26	- 17	- 33	- 61	- 15	- 18	- 20	- 7	- 6
Municipal Governments	- 4	- 3	3	- 7	2	1	- 3	-	3
Corporations	14	- 1	13	- 13	- 2	- 28	- 3	- 15	- 23
Total	-173	- 85	- 86	-164	- 36	- 60	- 6	24	- 7
U.S.:									
U.S. Government	- 1	- 1	1	- 1	2	- 11	5	- 12	3
Other	1	10	5	3	1	- 1	6	- 3	8
Total	-	9	6	1	3	- 12	10	- 15	11
Total bonds and debentures	-173	- 76	- 80	-163	- 33	- 72	5	9	4
Stocks									
Canadian	65	6	85	95	76	9	68	76	54
U.S.	7	10	8	28	22	44	30	- 2	8
Total stocks	72	15	92	122	98	53	98	75	62

1. Direct and guaranteed.

Source: The Dominion Bureau of Statistics

deficient, however, in one other major respect. Although they are broken down according to the type of borrower, no distinction is made between securities payable in Canadian dollars and those payable in U.S. dollars. This data limitation not only prevents a more detailed analysis of the net trade, but it also stands in the way of meaningful interpretation of other types of capital flow; this was discovered in the final parts of the section on new issues to be a result of the strong likelihood that there exists a high degree of substitution on the part of U.S. investors between new and outstanding Canadian issues of the same type. It does not, however, impede the regression analysis to follow, for the direction of response to the various independent variables would be the same regardless of the currency of repayment.

One other complication which did not arise in the case of new issues must be mentioned. This is the possibility of trade between the U.S. and third countries in Canadian securities (or trade between Canada and third countries in U.S. securities). It is possible that, with unchanged conditions in the United Kingdom, a U.S. investor will respond to an increase in the attractiveness of Canadian securities relative to U.S. securities by purchasing the former from a U.K. owner. This trade is not covered in the statistics used here, which apply only to U.S.-Canadian transactions. Its relative importance is not, however, likely to be very great.

Regressions were run using as the dependent variables the monthly net trade in outstanding securities of each of the various types from 1952-1960.[1] The results are found in Table 17, which is constructed after the model of Table 12.

It will be noticed that the net trade in every type of Canadian bond (and in total Canadian bonds), except for those of the Government of Canada, shows a significant slope coefficient associated with the interest rate differential. Indeed, the only problem in the cases of provincial and corporate bonds and debentures is the selection of the "best" interest rate differential variable from the several possibilities. The only interest rate variable bearing no apparent relation to net trade is that referring to municipal securities. Net trade in U.S. Governments and total U.S. bonds and debentures also demonstrates response to the federal government yield differential. As was shown above, these positive responses may be due either to long-run capital movements in search of income or to short-term flows seeking maximum (purely domestic) capital appreciation. Net trade in stocks could not be proven to be responsive to international differentials in stock yields, whether common or preferred.

Significant slope coefficients were also found for g_S, the proxy for short-term exchange rate expectations, in the net trade of every type of Canadian security except for corporate bonds, and in that of U.S. Government securities. These results indicate that much of the trade in outstanding securities is indeed of a short-term nature, taking place so as to reap

1. Strictly, from February 1952 until November 1960.

TABLE 17

Regression Results—Dependent Variables: Canadian Sales Minus Purchases of Outstanding Securities, by Types, Monthly, 1952–1960
(Thousands of dollars)

	Constant	$R_c - R_u$ $H_o: b_1 = 0$ $H_1: b_1 = (+)$	g_S $H_o: b_2 = 0$ $H_1: b_2 = (-)$	g_L $H_o: b_3 = 0$ $H_1: b_3 = (+)$	R^2
Bonds and Debentures					
Canadian:					
Govt. of Canada[1]					
Federal Govt. 10 year	+ 67,808 (50,572)	+ 2,932 (2,456)	- 4,999*** (979)	- 71,534 (51,678)	.230
Provincial Govts.[1]					
Federal Govt. 10 year	- 16,855 (18,843)	+ 2,887*** (915) B = 33.9	- 653** (365) B = 19.0	+ 13,792 (19,255)	.090
Prov. (State) Govt. 10 year[2]	- 47,981 (20,752)	+ 2,849*** (690) B = 44.4	- 568** (332) B = 16.5	+ 41,941** (20,677) B = 20.9	.144
Industrial 25 year[2]	- 57,048 (23,467)	+ 5,451*** (1,534) B = 41.2	- 400 (330)	+ 49,120** (22,818) B = 24.5	.112
Public Utility 25 year[2]	- 23,576 (22,372)	+ 2,423* (1,485)	- 283 (349)	+ 20,019 (22,272)	.027
Municipal Govts.[1]					
Federal Govt. 10 year	- 6,994 (5,329)	+ 428** (259) B = 23.0	- 163* (103) B = 17.2	+ 6,926 (5,445)	.043
Corporations					
Federal Govt. 10 year	- 51,337 (15,929)	+ 995* (774) B = 13.6	+ 313 (308)	+ 51,263*** (16,277) B = 29.8	.119
Industrial 25 year[2]	- 63,980 (20,061)	+ 1,754* (1,311) B = 15.4	+ 406 (282)	+ 62,371*** (19,507) B = 36.2	.120
Public Utility 25 year[2]	- 68,166 (17,997)	+ 2,675** (1,194) B = 24.0	+ 321 (281)	+ 66,562*** (17,917) B = 38.7	.146

1. Direct and guaranteed.
2. Term to maturity is approximate.

Table 17 (continued)

	Constant	$R_c - R_u$	g_S	g_L	R^2
Total Canadian					
Federal Govt. 10 year	- 60,314 (71,189)	+ 7,649** (3,458) B = 23.3	- 5,270*** (1,378) B = 39.7	+ 54,105 (72,745)	.128
U.S.:					
U.S. Govt.					
Federal Govt. 10 year	- 29,603 (13,999)	+ 1,070* (680) B = 17.2	- 387* (271) B = 15.4	+ 29,787** (14,305) B = 20.3	.059
Other U.S.					
No significant results					
Total U.S.					
Federal Govt. 10 year	- 32,287 (16,649)	+ 1,327* (809) B = 18.0	- 339 (322)	+ 32,607** (17,013) B = 18.8	.051
Total Bonds and Debentures					
Federal Govt. 10 year	- 46,493 (67,091)	+ 9,036*** (3,259) B = 28.5	- 5,811*** (1,299) B = 45.4	+ 39,824 (68,558)	.169
Stocks					
Canadian:					
Common Stock	- 40,725 (48,194)	- 3,188 (2,675)	- 2,380*** (782)	+ 49,034 (49,023)	.120
Preferred Stock	- 75,109 (38,955)	+ 141 (1,764)	- 1,954** (890) B = 26.9	+ 83,330** (40,130) B = 19.6	.107
U.S.:					
No significant results					
Total Stocks					
Common Stock	- 10,193 (61,552)	- 2,282 (3,416)	- 1,810** (999)	+ 18,516 (62,611)	.039

*** Significantly different from zero at the 1 per cent level (one-tailed test).
** " " " " " " 5 " " " " "
* " " " " " " 10 " " " " "

returns from expected changes in the exchange rate. The fact that the trade in Canadian Government bonds and Canadian stocks, two of the three high-volume types of trade, exhibit strong response to this factor is of particular interest. The total lack of response on the part of the net trade in outstanding Canadian corporate bonds suggests, as Radford surmised, [1] that transactions in this type of security reflect longer-run investment decisions; the same may be true of the net trade in U.S. stocks.

The spot exchange rate variable was expressed as the price of the U.S. dollar in terms of Canadian dollars. One would therefore expect the sign of the slope coefficient to be positive if the earlier reasoning with respect to this variable is correct. Significantly positive slope coefficients were, in fact, found in the cases of net trade in provincial and corporate (and, very nearly, municipal) bonds, and in U.S. Government and total U.S. bonds. Trade in Canadian stocks also demonstrated response to this factor in one of the computed regressions. The insignificant, indeed negative, coefficient associated with the spot rate in the case of Government of Canada securities may be the result of the same sort of reverse causation which made the use of this variable inappropriate in the case of new issues. On the other hand, it may imply that it is only short-term exchange rate expectations which are relevant to this type of trade. In opposite fashion, the net trade in corporate bonds and debentures is seen to be affected by exchange rate expectations after all, but only to those of a somewhat longer-run type.

C. Summary and Conclusions

Canadian provincial government, municipal government, and corporate borrowers all are influenced by Canadian–U.S. interest rate differentials in their decisions as to the currency in which new security issues are to be payable. This behavior characteristic of the greatest component of long-term capital flows between the U.S. and Canada is undoubtedly a major contributing factor to the close links which were earlier found to exist between equivalent U.S. and Canadian interest rates. It has not been possible, however, to establish conclusively that Canadian borrowers react to interest rate differentials of the types which would apply to their own issues. Rather, it appears that they are governed primarily by federal government yields, which probably most accurately portray the over-all state of capital markets in the two countries. The thinness of the Canadian markets for the issues of provincial and municipal governments and corporations make this a not unreasonable practice. (This practice, incidentally, is more accurately thought of as the consideration of over-all capital market conditions rather than the more narrow interpretation—consideration of federal government yields—suggested by the statistics.) With the exception of provincial governments, exchange rate expectations could not be proven to be of significant importance to Canadian borrowers, nor should they have been expected to play major roles in their security-issuing decisions.

---

1. R. A. Radford, op. cit., p. 238.

U.S. investors purchasing new issues (net) of Canadian securities payable in Canadian dollars also apparently respond to interest rate differentials where provincial, municipal, and, especially, corporate securities are concerned, but they do not do so with respect to their purchases of stocks. Yield differentials other than those referring to federal government bonds are found to be of greater importance to them than they are to Canadian borrowers. This is probably because individual investors, unless dealing in very large amounts, need not be as concerned as must borrowers with the implications of thin security markets. Investor practice in the case of Government of Canada new issues is unclear because of an apparent lack of U.S. interest in them before 1956. There is, in any case, some question as to the usefulness of analysis of U.S. investor behavior which ignores the simultaneously occurring trade in outstanding issues. The same is true of Canadian participation (net) in the U.S. new issues market, where no significant response to interest rate differentials was uncovered in the data. In neither of these investor-motivated types of capital flow were exchange rate expectations found to be of any import; this suggests that investments of this type were of a long-run nature.

Trade in outstanding bonds and debentures, since it clearly responds to Canadian–U.S. interest rate differentials, provides another major link between the capital markets of the two countries. This link appears to exert itself mainly through the trade in Canadian provincial, municipal, and corporate securities and that in U.S. Government bonds. Interest rate differentials are of greater importance (as shown by the significance of the slope coefficients and the relative sizes of the beta coefficients) than exchange rate expectations in the determination of net trade. The most significant of these interest rate responses are again found in the cases of non-federal government yield differentials, since these capital flows are also investor-motivated rather than borrower-motivated. As far as Government of Canada securities are concerned, the major determinant of net trade is not the relevant interest rate differential, but rather the state of expectations as to the future level of the exchange rate. This trade clearly must be comprised very largely of short-term capital flows. Net trade in common and preferred stocks responds to these exchange rate expectations but is not influenced by relative stock yields.

Thus, capital movements obviously do tend to bind the Canadian and U.S. capital markets. All types of Canadian issuers of U.S. dollar bonds and debentures have played major roles in this tying together of rates. A similar but much smaller role has been played by U.S. investors in the Canadian new issue market. However, both of these types of market-tying capital flows occur only at irregular intervals. It is the function of the trade in outstanding securities to furnish a steady stream of this tying capital. It should be remembered, however, that the trade in outstanding securities is not the only source of "steady" ties between the two markets. Among other major ones are the imitative forces which produce similar capital flows between them and the constant international flow of short-term capital of other types.

Unifying influences have therefore apparently played a major role in North American capital markets, a role which has been little impeded by the existence of factors (discussed in Section II of Part A) which would have permitted some divergences in observed interest rate behavior. If interest rates are considered to be at all significant as decision variables for monetary policy, the strength of these ties suggests the existence of problems for the Canadian monetary authorities.

The small size of the Canadian capital market relative to its U.S. counterpart makes it probable that the observed tendency towards unification of North American capital markets will exert itself through the adjustment of Canadian towards U.S. interest rates rather than vice versa. As a result, it would appear that the ability of the Bank of Canada to pursue independent interest rate policies was, and is, severely limited. This is not to state, however, that its ability to pursue a monetary policy independent of that of the U.S. authorities is so limited. The flexible Canadian exchange rate permits a substantial degree of independence for Canadian monetary policy even though very little for Canadian interest rates. The strength of the ties between U.S. and Canadian capital markets is such that international capital flows will be produced by any divergence of Canadian and U.S. interest rates from their normal relationship to one another. These capital flows will result in changes in the exchange rate which will have the desired effect upon the Canadian economy. For example, if Canadian monetary policy were to be tight relatively to U.S. monetary policy, there would be a tendency for more capital to flow from the U.S. into Canada (or less to flow from Canada into the U.S.), with the result that the value of the Canadian dollar would rise. This change in the exchange rate would put pressure upon Canadian export and import-competing industries. Conversely, a relatively easy Canadian policy would stimulate these industries. Thus, although it is likely that Canadian monetary policy which diverges significantly from concurrent U.S. policy would have a differential impact upon various sectors of the economy, there is no reason for it not to be effective.

As long as Canadian monetary policy remained "in step" with U.S. policy it would operate, of course, in the conventional manner through the effects of interest rates upon investment decisions. The factors discussed in Section II of Part A would permit a very limited amount of independence of interest rate behavior; the conventional effects of interest rates upon investment decisions could therefore still play a small role even in an independent Canadian monetary policy. In general, however, monetary independence is achieved through the existence of a flexible exchange rate.

Summary: This paper introduces a two-sector growth model which is designed to analyze the various patterns of growth and the evolution of the price system. The critical assumption is made that technological change is embodied in the "latest" model of capital goods. This gives rise to a whole spectrum of capital goods which differ in their productivity, and, consequently, in their relative prices.

The paper states the sectoral composition of the capital stock when the economy is in a competitive equilibrium, and the stability properties of the model. First, the assumptions are stated, and then the model itself is constructed. A critical concept of "terminal path" is introduced; it describes a path of growth in which all variables change at a constant proportional rate. Next, a special case is considered in which technical change follows a certain pattern. In this case, the behavior of the system when it is not on a terminal path is analyzed. Next, the general terminal path solutions are investigated and some of their interesting features demonstrated. Finally, the evolution of the relative prices of capital goods and the interest rate are analyzed.

It is assumed that the production functions in the two sectors are of the Cobb-Douglas type, and technical change is neutral; this makes the paper a complete extension of Solow's recent one-sector model.

This essay was submitted to Yale University in 1962 in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The author is currently a research associate at the Institute for Mathematical Studies in the Social Sciences at Stanford University.

Dissertation Committee: Professors Arthur M. Okun (Chairman), William J. Fellner, and Edmund S. Phelps.

*

I wish to thank Professor Arthur Okun, chairman of my dissertation committee, for his helpful suggestions and constant encouragement. Professor Edmund Phelps, the second member of my committee, made many valuable comments and I am indebted to him for his encouragement and interest in this work. I benefited also from penetrating observations made by Professor William Fellner, the third member of the committee.

My colleagues in the Economics Department at Yale University made many important comments on earlier drafts of this dissertation. I am grateful to Messrs. T. N. Srinivasan and William Brainard, with whom I had many stimulating discussions, for their penetrating suggestions. I am also indebted to Professor Edwin Mills of John Hopkins University for his very useful last-minute suggestions.

I am grateful to the Economics Department of Yale University for its financial support of my work.

Finally, I am indebted to Miss Nancy Bowles for stylistic improvements and to Miss Mary McNulty for excellent typing of both the early and the final draft.

M. K.

PATTERNS OF GROWTH AND VALUATION IN A TWO-SECTOR MODEL

Mordecai Kurz

Introduction

Modern developments of economic theory brought about impressive advances in the theory of economic growth. The great variety of models constructed for the purpose of analyzing the developing economy stem from great differences in tools of analysis, basic economic concepts and outlook, and assumptions regarding market structures.

Most of the modern "Neoclassical" growth models utilize the idea of an aggregate production function and a one-commodity economy. For many purposes, these models are sufficient, and a number of impressive results have been obtained with their aid. There are, however, many problems in growth theory that require at least a two-commodity model for their analysis.

(1) The Need for Two-Sector Growth Models

Two of the basic facts of the economic life of a growing economy are that it needs capital goods to maintain its growth and that these capital goods are produced within the system.

In reality, there are production units which produce consumer goods and production units which produce capital goods. In this sense, a two-sector model should be a better approximation to reality than a one-sector model. The analytical complexity of a two-sector model, however, is obviously greater than that of a one-sector model. This is due both to the presence of a more complex technology and to the existence of a more complicated set of necessary economic decisions.

The fact that capital goods and consumer goods are produced under different sets of technological conditions means that the growth of the economy depends upon both the technological conditions governing the production of consumer goods and those governing the production of capital goods. Clearly, a question concerning, for example, the effect on the growth process of the differences between the technological conditions governing the production of capital goods and those governing consumer goods cannot arise in a one-sector model; it can be answered only with the aid of a two-sector model.

The analysis of economic questions, as well as technological considerations, is more interesting and more complex when a two-sector model is used. We may be interested in the conditions that the two sectors must satisfy in order for competitive equilibrium to exist, e.g., what should be the allocation of the capital stock between the sectors? What should be the composition of new investments going into each sector? We may ask what is the effect of capital mobility between the two sectors? What is the effect of technical progress in one or both sectors?

One of the most interesting problems which cannot arise in a one-sector model but does arise in a two-sector model is as to the behavior of the price of capital relative to that of consumer goods. We are interested

in the factors which govern the level and the rate of change of this relative price. In addition, the discussion of the interest rate in its relations to the rate of growth of income and the rate of change of the price of capital becomes richer than in one-sector models.

There are instances in which conclusions drawn from a one-commodity model will not be valid in a model with two commodities. As an example, in those one-sector models in which the marginal productivity of capital cannot be negative and the interest rate equals the marginal productivity of capital, it is clear that the interest rate must be non-negative. This result does not necessarily hold for a two-sector model, even if the marginal productivity of capital in both sectors is non-negative.

(2) Technical Progress

One of the difficult questions faced in the construction of growth models is how to incorporate technical progress into the model. In a recent paper, R. M. Solow suggested a new way of incorporating technical progress.[1] He restricts himself to a Cobb-Douglas production function. Capital goods at the moment of their construction embody all the latest technical knowledge but do not change technologically thereafter. Associated with those capital goods produced at any particular point of time, there exists a distinct production function. We can therefore talk of different vintages of capital, but in general we shall use the expression "vintage v" to describe both the type of capital goods and the production function associated with vintage v capital. $K(v, t)$ denotes the amount of capital goods embodying the technological knowledge of time point v that are being used at time t. $L(v, t)$ is the labor input associated with the v^{th} vintage of capital. The change in the production function through time is assumed to be neutral; consequently, total output at time t can be written

$$F(t) = \int_{-\infty}^{t} Ae^{\lambda v} K(v, t)^{1-\alpha} L(v, t)^{\alpha} dv \quad . \tag{1}$$

Here, $e^{\lambda v}$ is the technical improvement factor associated with vintage v capital. Interpreting capital goods as "machines," $K(t, t)$ may be regarded as the number of new machines added to the "stock of machines" at time t. Given the assumption of profit maximization, the optimal distribution of labor over the vintages is determined by equating the wage rate to the marginal productivity of labor in all uses. With these assumptions, Solow has shown that total output may be written in the following way:

$$F(t) = AL(t)^{\alpha} K(t)^{1-\alpha} \quad , \tag{2}$$

---

1. R. M. Solow, "Investment and Technical Progress," in Mathematical Methods in the Social Sciences 1959, Arrow, Karlin, and Suppes, eds., (Stanford: Stanford University Press, 1960), pp. 89-104.

where $K(t) = \int_{-\infty}^{t} e^{\sigma v} K(v, t)dv$, $\sigma = \frac{\lambda}{1-\alpha}$ and $L(t)$ is the total labor employed in the economy. Note that $K(t)$ is a weighted sum of the amounts of capital goods of the various vintages, the weights being given by $e^{\lambda v}$. Although the stock of capital is heterogeneous, expression (2) suggests that we may look at $K(t)$ as "effective capital." Solow has derived some of the properties of this model concerning the valuation of assets and obsolescence, but he recognizes the fact that these problems are "more naturally posed in a two-sector model."[1]

(3) Plan of Work

Our plan in this work is to extend Solow's model into a general equilibrium, two-sector growth model of the economy. The model will describe various patterns of growth that the economy is capable of exhibiting. We are interested in the conditions that such an economy must satisfy in order for competitive growth equilibrium to prevail. These conditions will also specify the relationships between corresponding variables in the two sectors. In this connection, we shall analyze the sectoral composition of the capital stock, total labor input, and new investments. We shall also be concerned with the evolution of the capital-labor ratio in each sector and the relationships between the two ratios.

It will be assumed that the ratio of total savings to total income is a constant. We shall be concerned with the influence of this savings ratio on the level and the rate of change of the economic variables. This question has particular importance for public policy directed towards affecting the rate of growth of the economy.

The next group of questions to be analyzed concern the price implications of competitive growth. We are interested in the determinants of the price of capital relative to consumer goods and the behavior of this price under various patterns of growth. We want to be able to state the conditions under which this price will rise, fall, or remain constant through time. Since the capital stock in the model is heterogeneous, we shall analyze the behavior of both the price of the "latest" model of capital goods and the price of "old machines." As technical change is introduced, these "old machines" become obsolete and gradually lose their value. We shall be concerned both with the factors which govern the rate of obsolescence and with the intensity of this process.

Finally, the interest rate as defined in the model will be a subject for our investigation. We are concerned with the relationships among the interest rate, the productivity of capital, the rate of growth of income, and the rates of change of the prices of capital goods. We wish to know under what conditions the interest rate can become negative.

---

1. Ibid., p. 102.

I. The Construction of the Model

In the introduction, we have indicated the basic concepts that we intend to use in our model. Let us now proceed to construct the model in a rigorous fashion.

(1) Assumptions of the Model

(a) There are two sectors: a consumer goods sector and a capital goods sector.

(b) New capital goods always embody the latest technical improvements. Once these goods are installed, no further improvements are possible. A production function is associated with each vintage of capital goods, and capital goods of a given vintage can only be used with their respective production functions.

(c) Before installation, a new capital good can be placed in either sector, but having been installed in one sector, it has to stay there until the end of its economic life, i.e., it cannot be shifted to the other sector. We shall refer to this assumption as the "non-shiftability assumption."

(d) Labor is exogeneously and inelastically supplied to the market. The supply is assumed to be described by the exponential function $L(t) = L(0)e^{nt}$ where n is the constant percentage rate of growth of the labor force.

(e) Capital goods do not depreciate physically.

(f) The production functions of the two sectors are assumed to be of the Cobb-Douglas type with a neutral technological improvement factor. In symbols, if $F(v, t)$ and $G(v, t)$ are the output of consumer goods and capital goods (respectively) in the v^{th} vintage of the respective sector, then

$$F(v, t) = Be^{\lambda v}L(v, t)^{\alpha}K(v, t)^{1-\alpha}$$

and

$$G(v, t) = Ae^{\gamma v}L_g(v, t)^{\beta}K_g(v, t)^{1-\beta} ,$$

where $L(v, t)$ and $L_g(v, t)$ are the respective labor inputs and $K(v, t)$ and $K_g(v, t)$ are the respective capital inputs. Constant returns to scale are assumed.

(g) The savings ratio is a constant, so that $S(t) = sY(t)$ is the savings function.

(h) Perfect competition prevails; labor is free to move within and between sectors and all factors are being paid the value of their marginal product.

(i) There exists no uncertainty concerning the future. Alternatively, all expectations are correct and they do materialize.

As a matter of convenience, a unit of consumer goods at each time t is used as a numéraire; the price of consumer goods is identically 1 at all points of time. All prices of capital goods are expressed in terms of consumer goods.

We use the terms "labor force" and "population" interchangeably since we assume that the total population is employed.

(2) Notation

In developing our model the following notation will be used:

$F(v, t)$ — amount of consumer goods produced by vintage v at time t.

$G(v, t)$ — amount of capital goods produced by vintage v at time t.

$L(v, t)$ — amount of labor employed by vintage v of consumer goods sector at time t.

$L_g(v, t)$ — amount of labor employed by vintage v of capital goods sector at time t.

$K(v, t)$ — amount of capital of vintage v employed in the consumer goods sector at time t.

$K_g(v, t)$ — amount of capital of vintage v employed in the capital goods sector at time t.

$r(v, t)$ — capital-labor ratio of factors employed in vintage v of the consumer goods sector at time t.

$r_g(v, t)$ — capital-labor ratio of factors employed in vintage v of the capital goods sector at time t.

$R(v, t)$ — rent paid per unit of capital of vintage v employed in the consumer goods sector at time t.

$R_g(v, t)$ — rent paid per unit of capital of vintage v employed in the capital goods sector at time t.

$a(v, t)$ — price of capital goods of vintage v employed in the consumer goods sector at time t.

$a_g(v, t)$ — price of capital goods of vintage v employed in the capital goods sector at time t.

$F(t)$ — total output of consumer goods at time t.

$G(t)$ — total output of new capital goods at time t.

$q(t)$ — the price of a unit of new capital goods at time t.

$G_q(t)$ — value of output of newly built capital goods at time t.

$L(t)$ — total labor force at time t.

$l(t)$ — total labor employed in the consumer goods sector at time t.

$l_g(t)$ — total labor employed in the capital goods sector at time t.

$W(t)$ — the wage rate at time t.

$i(t)$ — the interest rate at time t.

$Y(t)$ — total income at time t.

$S(t)$ — total flow of savings (value) at time t.

$I(t)$ — income per capita at time t.

$C(t)$ — consumption per capita at time t.

(3) The Construction of the Model

(3.a) The General Formulation

Let us first examine the general form of the model without specifying the particular form of the production function (assumption f).

We define income to be the total value of output of the two sectors, consumer goods being taken as numéraire. We then have

$$Y(t) = \int_{-\infty}^{t} F(v, t)dv + q(t) \int_{-\infty}^{t} G(v, t)dv \ . \tag{1}$$

It follows from the assumptions of competitive equilibrium and mobility of labor that the value of the marginal product of labor employed at each vintage is equal to the wage rate. We then have

$$\frac{\partial F(v, t)}{\partial L(v, t)} = W(t) \text{ for all } v \tag{2}$$

and

$$q(t) \frac{\partial G(v, t)}{\partial L_g(v, t)} = W(t) \text{ for all } v. \tag{3}$$

Strictly speaking, we should write $v \leq t$ in (2) and (3), but since v is the date of construction while t is the date of employment, this inequality is understood always to hold.

Total employment is equal to

$$L(0)e^{nt} = \int_{-\infty}^{t} L(v, t)dv + \int_{-\infty}^{t} L_g(v, t)dv \ . \tag{4}$$

From assumption (g), total savings are related to income by the equation

$$S(t) = sY(t) \ . \tag{5}$$

However, the total amount saved is equal to the value of the output of the capital goods sector; hence,

$$S(t) = q(t) \int_{-\infty}^{t} G(v, t)dv \ . \tag{6}$$

$q(t)$ is the price per unit of newly built capital goods while $\int_{-\infty}^{t} G(v, t)dv$ is the total output of new capital goods. This total has to satisfy one more condition:

$$\int_{-\infty}^{t} G(v, t)dv = K(t, t) + K_g(t, t) \quad . \tag{7}$$

(7) simply says that the total output of the capital goods sector at each point in time is allocated back in the form of new investments, $K(t, t)$ in the consumer goods sector and $K_g(t, t)$ in the capital goods sector. We are left with the valuation side. During the life of each piece of equipment it will pay rental. In our model, there is no physical depreciation, but due to economic reasons a capital good may lose its value and eventually be withdrawn from service.[1] When a capital good is installed, however, the investor estimates the future rentals that he will receive. By assumption (i), the investor knows all his future rentals, and by assumption (h), perfect competition prevails. Therefore, in equilibrium, the cost of production of a unit of new capital goods equals the present value of future rentals in either sector, the present value being computed with a common interest rate. Consequently

$$q(t) = \int_t^{\infty} \frac{\partial F(t, z)}{\partial K(t, z)} e^{-\int_t^z i(u)du} dz \tag{8}$$

and

$$q(t) = \int_t^{\infty} \frac{\partial G(t, z)}{\partial K(t, z)} q(z) e^{-\int_t^z i(u)du} dz \quad . \tag{9}$$

Equations (8) and (9) are based on the assumptions stated above. They could, however, also be viewed as the definitions of the uniform interest rate. Note that if we differentiate (8) and (9) with respect to t, we obtain[2]

$$\dot{q}(t) = \frac{\partial F(t, t)}{\partial K(t, t)} + \int_t^{\infty} \frac{d}{dt}\left[\frac{\partial F(t, z)}{\partial K(t, z)}\right] e^{-\int_t^z i(u)du} dz + i(t)q(t) \tag{8a}$$

and

$$\dot{q}(t) = -q(t) \frac{\partial G(t, t)}{\partial K_g(t, t)} + \int_t^{\infty} q(t) \frac{d}{dt}\left[\frac{\partial G(t, z)}{\partial K_g(t, t)}\right] e^{-\int_t^z i(u)du} dz + i(t)q(t) \quad . \tag{9a}$$

To interpret (8a) and (9a), we have to realize the meaning of $q(t)$. It is the price of a unit of the latest vintage of capital goods, where the word "latest"

---

1. We shall see that under assumption (f) a capital good is never withdrawn from service.

2. A dot on top of a variable designates its time derivative.

is understood in a dynamic sense. Hence $\dot{q}(t)$ measures the difference between the price of a unit of vintage $t + dt$ capital at $t + dt$, and the price of a unit of vintage t capital at t. (8a) and (9a) say that the change in price is equal to the interest charges (third term) minus the value of the marginal productivity of capital (first term) plus the discounted value of the increase in the marginal productivity of the new capital at t (second term) due to technological progress at t. It should be noted that the presence of the second term is a direct consequence of the assumption that technical progress is embodied in the "latest" model.

(3.b) Specialization of Assumption (f)
By definition (see (1) above)

$$Y(t) = \int_{-\infty}^{t} Be^{\lambda v} L(v,t)^{\alpha} K(v,t)^{1-\alpha} dv + q(t) \int_{-\infty}^{t} Ae^{\gamma v} L_g(v,t)^{\beta} K_g(v,t)^{1-\beta} dv \ . \qquad (10)$$

By assumption (h), all vintages pay the same wage rate; hence, equating the value of marginal productivity to the wage rate means (for the consumer goods sector)

$$\alpha Be^{\lambda v} L(v,t)^{\alpha-1} K(v,t)^{1-\alpha} = W(t) \text{ for all } v \ . \qquad (11)$$

Since (11) is true also for $v = t$ we have

$$\alpha Be^{\lambda v} L(v,t)^{\alpha-1} K(v,t)^{1-\alpha} = \alpha Be^{\lambda t} L(t,t)^{\alpha-1} K(t,t)^{1-\alpha} \ .$$

Dividing both sides by αB and using the notation of the capital-labor ratio $r(v,t)$, we obtain

$$r(v,t) = r(t,t)e^{\frac{\lambda}{1-\alpha}(t-v)} \qquad (12)$$

and equivalently

$$r(v,t)e^{\frac{\lambda}{1-\alpha}v} = r(t,t)e^{\frac{\lambda}{1-\alpha}t} \text{ for all } v \ . \qquad (12a)$$

This equation says that when labor is distributed optimally over the vintages, then the ratio between the capital-labor ratio in the t^{th} and the v^{th} vintages is inversely related to the difference $(t-v)$ as shown in the diagram below, which depicts equation (12). We then see that the older a vintage of capital is, the higher is the ratio of capital to labor in that vintage. Note also (equation (12a)) that the product $r(v,t)e^{\frac{\lambda}{1-\alpha}v}$ is a constant for each t, independent of v. Now, since by definition $r(v,t) = \frac{K(v,t)}{L(v,t)}$, we have that

$$\frac{K(v,t)e^{\frac{\lambda}{1-\alpha}v}}{L(v,t)} = \frac{K(t,t)e^{\frac{\lambda}{1-\alpha}t}}{L(t,t)} \text{ independently of } v \ . \text{ If we regard } K(v,t)e^{\frac{\lambda}{1-\alpha}v}$$

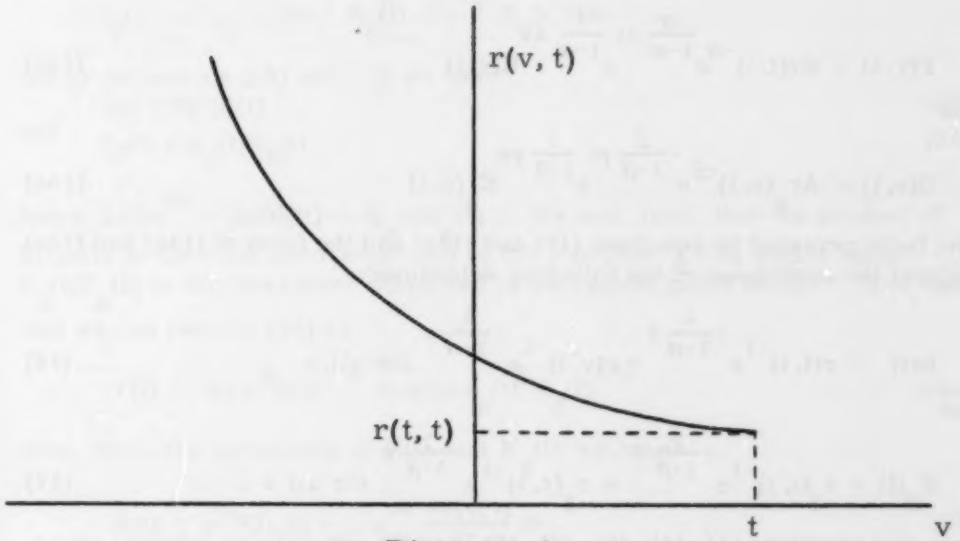

Diagram 1

as "effective capital," we may say that the ratio between "effective capital" and labor is the same for all vintages.[1] A similar argument holds for the capital goods sector. Let us keep these considerations in mind for further use later.

By rewriting (12) we can obtain labor employment:

$$L(v, t) = \frac{K(v, t)}{r(t, t)} e^{\frac{\lambda}{1-\alpha}(v-t)} \tag{13a}$$

and

$$L_g(v, t) = \frac{K_g(v, t)}{r_g(t, t)} e^{\frac{\gamma}{1-\beta}(v-t)} \tag{13b}$$

We know from (10) that the outputs of the v^{th} vintage in the two sectors are

$$F(v, t) = Be^{\lambda v} L(v, t)^{\alpha} K(v, t)^{1-\alpha} \tag{14}$$

and

$$G(v, t) = Ae^{\gamma v} L_g(v, t)^{\beta} K_g(v, t)^{1-\beta} \tag{15}$$

Using (13) we can write (14) and (15) as

1. This also raises the question of the units in which capital is measured. Discussion of this question follows below.

$$F(v,t) = Br(t,t)^{-\alpha} e^{\frac{\alpha}{1-\alpha}\lambda t} e^{\frac{1}{1-\alpha}\lambda v} K(v,t) \tag{14a}$$

and

$$G(v,t) = Ar_g(r,t)^{-\beta} e^{-\frac{\beta}{1-\beta}\gamma t} e^{\frac{1}{1-\beta}\gamma v} K_g(v,t) \ . \tag{15a}$$

The facts revealed by equations (12) and (12a) and the form of (14a) and (15a) suggest the usefulness of the following definitions:

$$R(t) = r(t,t)^{-1} e^{-\frac{\lambda}{1-\alpha}t} = r(v,t)^{-1} e^{-\frac{\lambda}{1-\alpha}v} \quad \text{for all } v \tag{16}$$

and

$$R_g(t) = r_g(t,t)^{-1} e^{-\frac{\gamma}{1-\beta}t} = r_g(v,t)^{-1} e^{-\frac{\gamma}{1-\beta}v} \quad \text{for all } v \ . \tag{17}$$

Note that equations (1), (4), (6), (7), (8), and (9) are difficult integral equations. We want to reformulate the model so as to avoid working with such equations. For mathematical convenience, we introduce the following definitions:

$$K(t) = \int_{-\infty}^{t} e^{\frac{1}{1-\alpha}\lambda v} K(v,t)dv \ , \tag{18}$$

$$K_g(t) = \int_{-\infty}^{t} e^{\frac{1}{1-\beta}\gamma v} K_g(v,t)dv \ , \tag{19}$$

$$\frac{\lambda}{1-\alpha} = \sigma \ , \tag{20}$$

and

$$\frac{\gamma}{1-\beta} = \sigma_g \ . \tag{21}$$

$K(t)$ can be viewed as the "effective capital stock" in the consumer goods sector and $K_g(t)$ as the corresponding variable in the capital goods sector. To obtain income, $Y(t)$, we integrate (14a) and (15a) over all vintages. By using (16)–(21) we obtain

$$Y(t) = BR(t)^{\alpha} K(t) + q(t)AR_g(t)^{\beta} K_g(t) \ , \tag{22}$$

which involves no more integrals. Note now that the labor supply = demand equation can be obtained as follows: from (13) we have

$$1(t) = \int_{-\infty}^{t} L(v,t)dv = R(t) \int_{-\infty}^{t} e^{\sigma v} K(v,t)dv$$

and

$$l_g(t) = \int_{-\infty}^{t} L_g(v,t)dv = R_g(t) \int_{-\infty}^{t} e^{\sigma_g v} K_g(v,t)dv$$

and by definitions (18) and (19) we have

$$l(t) = R(t)K(t) \tag{23}$$

and

$$l_g(t) = R_g(t)K_g(t) \ , \tag{24}$$

hence $L(0)e^{nt} = R(t)K(t) + R_g(t)K_g(t)$. We see, then, that the product of $R(t)K(t)$ is the total labor employed in the consumer goods sector while $R_g(t)K_g(t)$ is the total labor employed in the capital goods sector. This means that we can rewrite (22) as

$$Y(t) = Bl(t)^{\alpha}K(t)^{1-\alpha} + q(t)Al_g(t)^{\beta}K_g(t)^{1-\beta} \ . \tag{22a}$$

Now, from the definitions of $K(t)$ and $K_g(t)$ we have

$$\dot{K}(t) = e^{\sigma t}K(t,t) + \int_{-\infty}^{t} e^{\sigma v} \frac{\partial K(v,t)}{\partial t} dv \tag{25}$$

and

$$\dot{K}_g(t) = e^{\sigma_g t}K_g(t,t) + \int_{-\infty}^{t} e^{\sigma_g v} \frac{\partial K_g(v,t)}{\partial t} dv \ . \tag{26}$$

But, by the non-shiftability assumption,

$$\left.\frac{\partial K(x,y)}{\partial y}\right|_{\substack{x=v \\ y=t}} = \left.\frac{\partial K_g(x,y)}{\partial y}\right|_{\substack{x=v \\ y=t}} = 0 \text{ for all } v \ .$$

Here $K(x,y)$, the amount of capital of vintage x employed at time y in the consumer goods sector, is viewed as a function of two variables: the vintage and time. The above equation states that the quantity of capital of any vintage does not change after its installation. Now, the total volume of new capital goods, $K(t,t) + K_g(t,t)$, is obtained from (25)–(26):

$$K(t,t) + K_g(t,t) = \dot{K}(t)e^{-\sigma t} + \dot{K}_g(t)e^{-\sigma_g t} \ . \tag{27}$$

The wage rate equations can now be stated as

$$\alpha BR(t)^{\alpha-1} = W(t)$$

and

$$q(t)\beta AR_g(t)^{\beta-1} = W(t) \ .$$

The rentals on capital goods of vintage t employed in the consumer goods sector at all $z > t$ is $B(1-\alpha)e^{\sigma t}R(z)^{\alpha}$. If the instrument is employed in the capital goods sector, then its rent is $q(z)A(1-\beta)e^{\sigma_g t}R_g(z)^{\beta}$. Hence we have

$$q(t) = \int_t^{\infty} B(1-\alpha)e^{\sigma t}R(z)^{\alpha}e^{-\int_t^z i(u)du}\,dz$$

and

$$q(t) = \int_t^{\infty} q(z)A(1-\beta)e^{\sigma_g t}R_g(z)^{\beta}e^{-\int_t^z i(u)ju}\,dz\,.^{1}$$

It follows that

$$\frac{\dot{q}(t)}{q(t)} = -\frac{1}{q(t)}B(1-\alpha)e^{\sigma t}R(t)^{\alpha} + i(t) + \sigma$$

and

$$\frac{\dot{q}(t)}{q(t)} = -A(1-\beta)e^{\sigma_g t}R_g(t)^{\beta} + i(t) + \sigma_g\,.$$

By subtracting one from the other and transposing terms we have

$$\frac{1}{q(t)}B(1-\alpha)e^{\sigma t}R(t)^{\alpha} - A(1-\beta)e^{\sigma_g t}R_g(t)^{\beta} = \sigma - \sigma_g\,. \tag{28}$$

(28) will prove a very useful equation. This completes the construction. Let us summarize by setting forth the entire system:

$$Y(t) = BR(t)^{\alpha}K(t) + q(t)AR_g(t)^{\beta}K_g(t) \tag{29}$$

$$\alpha BR(t)^{\alpha-1} = W(t) \tag{30}$$

$$q(t)\beta AR_g(t)^{\beta-1} = W(t) \tag{31}$$

$$S(t) = sY(t) \tag{32}$$

$$S(t) = q(t)AR_g(t)^{\beta}K_g(t) \tag{33}$$

---

1. In writing the = sign in these two equations, we assume that new investments take place in both sectors. It is conceivable that new capital goods will be invested in one sector only for some time. It can be shown that this can continue for a finite time only. After this period, investment will take place in both sectors. Our analysis proceeds from this point on.

$$s(t) = q(t)\left[K(t,t) + K_g(t,t)\right] \tag{34}$$

$$K(t,t) = \dot{K}(t)e^{-\sigma t} \tag{35}$$

$$K_g(t,t) = \dot{K}_g(t)e^{-\sigma_g t} \tag{36}$$

$$L(0)e^{nt} = R(t)K(t) + R_g(t)K_g(t) \tag{37}$$

$$q(t) = \int_{-\infty}^{t} B(1-\alpha)e^{\sigma t}R(z)^{\alpha}e^{-\int_t^z i(u)du}\,dz \tag{38}$$

$$q(t) = \int_{-\infty}^{t} q(z)A(1-\beta)e^{\sigma_g t}R_g(z)^{\beta}e^{-\int_t^z i(u)du}\,dz \tag{39}$$

In the system (29)–(39), we replace (38) or (39) by equation (28). Together (29)–(37), (28) and either (38) or (39) form a system of 11 equations in the 11 unknown time functions: $Y(t)$, $S(t)$, $q(t)$, $W(t)$, $R(t)$, $K(t)$, $R_g(t)$, $K_g(t)$, $K(t,t)$, $K_g(t,t)$, and $i(t)$. Our aim is to solve and analyze this system.

(3.c) A Note on the Units of Capital Goods

The reader may wonder as to the units in which capital goods of, say, vintage v capital are measured. We shall answer this question now.

In our model, we are free to choose arbitrarily two units: the unit of the homogeneous consumer good and the unit for capital goods of some arbitrary vintage. Let us choose a unit for vintage 0 capital. We know that the production function associated with vintage 0 capital always produces the latest model of capital goods. At some point of time t we can write

$$G(0,t) = AL(0,t)^{\beta}K(0,t)^{1-\beta}.$$

If we take $\frac{1}{A}$ units of labor and $\frac{1}{A}$ units of capital goods of vintage 0, our production function will produce one unit of capital of vintage t for all t. Indeed, we see that if these inputs are kept constant through time, the above production function will generate one unit of capital at every point in time. The unit of capital goods does not change through time; only the quality of capital goods changes in the way we have assumed it to change.

(4) Some Remarks on the Solution of the Model

(4.a) The Method of Solution

It can be shown that the 11-equation system can be reduced by a process of elimination to the following 4-equation system:[1]

1. The sub-system (40a)–(40b) has the special property that its unknowns are the "effective variables" of the original system ($K(t)$, $K_g(t)$, $R(t)$, and $R_g(t)$).

$$L(0)e^{nt} = R(t)K(t) + R_g(t)K_g(t) \tag{40a}$$

$$AK_g(t)R_g(t)^{\beta} = \dot{K}(t)e^{-\sigma t} + \dot{K}_g(t)e^{-\sigma_g t} \tag{40b}$$

$$R(t)K(t) = \frac{(1-s)\alpha}{s\ \beta} K_g(t)R_g(t) \tag{40c}$$

$$\frac{\beta(1-\alpha)}{\alpha(1-\beta)}\ e^{\sigma t}R(t) = e^{\sigma_g t}R_g(t) + \frac{(\sigma-\sigma_g)}{A(1-\beta)}\ R_g(t)^{1-\beta} \tag{40d}$$

It is clear, then, that further reduction of the system will bring it to a single differential equation which we shall call "the basic differential equation." Even after the solution for this equation is obtained for each one of the cases to be considered, there is still one more step to carry out: the variables $R(t)$, $K(t)$, $R_g(t)$, and $K_g(t)$ have to be translated back into the variables that they conceal, e.g., the capital-labor ratios.

(4.b) Stability Analysis[1]

In the presentation of the results, the following important concepts will be used.

Definition 1: A set of values of the 11 variables is said to be a competitive equilibrium if it satisfies equations (29)–(39). We shall use the shorter term "equilibrium" to mean "competitive equilibrium."[2]

We define the time point $t = 0$ to be the starting point for our analysis. The history of the system for $t < 0$ is regarded at time zero as the initial conditions. In obtaining the competitive solution for $t = 0$, we state the conditions that the initial conditions have to satisfy in order for competitive equilibrium to obtain and be maintained for all $t \geq 0$.

Definition 2: A collection of 11 time functions whose set of values for each t is a competitive equilibrium is said to be a terminal path if each

---

1. For an interesting discussion see Thomas F. Dernburg and James P. Quirk, "Per Capita Output and Technological Progress," Institute for Quantitative Research in Economics and Management, Purdue University, Institute Paper No. 7 (1960).

2. A solution for the model is compatible with the usual notion of competitive equilibrium except for the fact that our preference ordering is peculiar; it is defined on income and requires a constant savings ratio over time. In the absence of a satisfactory theory of choice over time we shall restrict ourselves to this assumption. For the usual definition of preference ordering and its relationship to the definition of competitive equilibrium, see for example K. J. Arrow and G. Debreu, "Existence of an Equilibrium for a Competitive Economy," Econometrica, XXII (1954), pp. 265–90.

one of the 11 time functions develops through time at a constant rate of change.

A "terminal path" is usually called a "golden age solution" in the economics literature. It is felt, however, that this is an unfortunate term.

In general, there are initial conditions for the system which will place it in competitive equilibrium;[1] however, these conditions are not necessarily sufficient for the system to be on a terminal path. Three questions arise immediately:

1. Does a terminal path exist?
2. If it exists, is it unique? (Uniqueness means that there is only one set of initial conditions which place the system on a terminal path.)
3. If it is unique, does the system, in the limit, approach it (in some well-defined sense)?

Stability analysis deals with the third of these questions. In answering this question we shall be concerned only with the variable "income per capita." If I(t) is any equilibrium solution for income per capita, and if a unique "terminal path" solution for I(t) exists, namely $\bar{I}(t)$, $(I(t) \neq \bar{I}(t))$, then we shall make the following

Definition: The system is said to be stable if $\lim_{t \to \infty} (I(t)-\bar{I}(t)) = 0$.[2]

---

1. We shall be concerned with the conditions for the existence of a competitive equilibrium.

2. Stability of growth processes have been considered many times in the economic literature. P. A. Samuelson and R. M. Solow ["Balanced Growth Under Constant Returns to Scale," Econometrica, XXI (July 1953), pp. 412-24] have considered an n commodity model with outputs $v_1 v_2 v_3 \ldots v_n$. The model was said to be stable if, starting from any initial conditions, the ratios $v_1 : v_2 : v_3 : \ldots : v_n$ converged to some ratios $v_1^* : v_2^* : \ldots v_n^*$. T. W. Swan ["Economic Growth and Capital Accumulation," The Economic Record, XXXII (November 1956), pp. 334-61] and H. Uzawa ["Neutral Inventions and the Stability of Growth Equilibrium," Review of Economic Studies, XXVIII (2) No. 76 (February 1961), pp. 117-24] define a stable one-sector model to be one in which the capital-output ratio converges to a constant. Solow ["A Contribution to the Theory of Economic Growth," Quarterly Journal of Economics, LXX (1956), pp. 65-94] defines stability to be the convergence to a constant of the capital-labor ratio, although the capital-labor ratio never converges when technical progress is present in the model. There is some similarity between our definition of stability and the Swan-Uzawa definition, although ours imposes stronger requirements. It can be shown that a model which is stable in our sense must also be stable in the Swan-Uzawa sense.

As an example of the general form in which solutions are presented in this work, we shall take income $Y(t)$. Then $Y(0)$ designates the initial level of income at time $t = 0$, $\bar{Y}(0)$ is the unique (if it exists) initial terminal path level of income, and $Y(t)$ will be shown as a function of $\bar{Y}(0)$, $Y(0) - \bar{Y}(0)$, and time. The same will be true for all other variables.

(4.c) The Problem of the "General" Solution

The "general" case is the one in which $\lambda > 0$, $\gamma > 0$, $n > 0$, $0 < \alpha < 1$, $0 < \beta < 1$, $0 < s < 1$, $A > 0$, and $B > 0$. In this case, both labor and capital are used by the two sectors, constant returns to scale prevail, technological progress takes place in both sectors, and consumers neither consume nor save all their income. Here we impose no restrictions on the relationships between these parameters; e.g., the rates of technical progress may or may not be the same in both sectors. Without going into details, we can say that a general solution is possible in principle. Unfortunately, the mathematical complexity of the solution is so great that it is hard to solve it and more complex to interpret it. Therefore, only "terminal path" solutions will be obtained. We cannot carry out stability analysis because, in these solutions, the initial conditions of the system must be consistent with the terminal path.

It is interesting, however, that if we put some a priori restrictions on the parameters, say $b = 1$ or $\sigma = \sigma_g$, and solve the system for these special cases, we can obtain most of the economic properties of the model. In the present paper, we shall restrict ourselves to the "usual" range of the parameters α and β, namely $0 < \alpha < 1$ and $0 < \beta < 1$, and also $\lambda > 0$ and $\gamma > 0$. This means that we shall not present here the behavior of the system on the boundary and when technical progress is absent.

(5) Preliminary Results

The existence of competitive equilibrium imposes upon our variables a certain path of behavior which depends upon the special case at hand. There are, however, some general relationships that are always true. Let us observe these now.

Following directly from the assumption of a constant savings ratio we have that

$$\frac{F(t)}{G_q(t)} = \frac{\text{Value of consumer goods produced at } t}{\text{Value of new capital goods produced at } t} = \frac{1-s}{s} .$$

Consequently, the rate of growth of $F(t)$ must always be equal to the rate of growth of $G_q(t)$, namely: $\frac{\dot{F}(t)}{F(t)} = \frac{\dot{G}_q(t)}{G_q(t)}$ for all t when competitive equilibrium prevails.

The competitive equilibrium conditions on the sectoral composition of the labor force are as follows:[1]

1. The interested reader can easily derive these conditions from the

$l(t)$ = Total labor employed in the consumer goods sector

$$= \frac{\alpha(1-s)}{\alpha(1-s)+\beta s} \; L(0)e^{nt} \tag{41}$$

$l_g(t)$ = Total labor employed in the capital goods sector

$$= \frac{\beta s}{\alpha(1-s)+\beta s} \; L(0)e^{nt} \; . \tag{42}$$

Obviously $\frac{l(t)}{l_g(t)} = \frac{\alpha(1-s)}{\beta s}$. That is, the ratio between labor inputs in the two sectors is a constant.

The conditions on relative factor shares are:

Relative labor share $= (1-s)\alpha + s\beta$

Relative capital share $= (1-s)(1-\alpha) + s(1-\beta)$.

This shows that in a two-sector model relative shares are sensitive to the savings ratio.

These last few results show also how the model builds its own "aggregation procedure." The savings and consumption ratios become the weights of the above sums. Note also that if $L(t)$ is the total supply of labor, then it follows from (1a) that

$$\frac{l(t)}{L(t)} + \frac{l_g(t)}{L(t)} = (1-s)\,\frac{\alpha}{\alpha(1-s)+\beta s} + s\,\frac{\beta}{\alpha(1-s)+\beta s} \; .$$

Obviously, the expression on either side is always equal to 1, but the form of the components is interesting.

It follows from the constancy of the factors' shares that the rate at which the wage rate rises must be equal to the rate at which income per capita grows. Consequently, a general fact is that

$$\frac{\dot{Y}(t)}{Y(t)} = n + \frac{\dot{W}(t)}{W(t)} \; ,$$

or, the rate of growth of income is equal to the rate of growth of the labor force plus the rate at which the wage rate rises.

In general, the existence of competitive equilibrium imposes a certain relationship between the capital stocks of the two sectors; that is, there exists a certain composition of capital which is optimal. If we consider, for example, the case of no technical change, then $K(t)$ is the capital stock (in the usual sense) in the consumer goods sector and $K_g(t)$ is the capital stock

---

sub-system (40) by using equations (40a) and (40c). One solves first for $R(t)K(t)$ and then for $R_g(t)K_g(t)$. From equation (23) and (24), we know that $l(t) = R(t)K(t)$ and $l_g(t) = R_g(t)K_g(t)$.

in the capital goods sector. The competitive equilibrium composition has to be

$$K(t) = \frac{(1-s)(1-\alpha)}{s(1-\beta)} K_g(t) . \tag{43}$$

The composition of the capital stock will be different for different cases, but the existence of some relationship is a general fact.

II. Technical Similarity: $\sigma = \sigma_g$

We indicated at the end of Section I the difficulties involved in obtaining the solutions for the general case in which the economy is not on the terminal path. We analyze our model, therefore, by considering special cases. Among those cases in which technical progress is present and for which we can obtain the solutions for non-terminal path situations, the case $\sigma = \sigma_g$ is the most general. In the present section we shall present the results for this case.

In the first part of this section, we shall give an intuitive explanation for σ and σ_g, and the significance of the assumption $\sigma = \sigma_g > 0$. The results will be described in the second part. The third part will contain the solutions in formulae, while some proofs and derivations will be presented in the fourth part.

(1) The Meaning of σ and σ_g

We start by inspecting $R(v,t)$ and $R_g(v,t)$, the rentals paid out at t to a unit of the v^{th} vintage capital employed in the consumer and the capital goods sectors respectively. We have that[1]

$$R(v,t) = B(1-\alpha)e^{\sigma v} R(t)^{\alpha} \tag{1}$$

and

$$R_g(v,t) = q(t)A(1-\beta)e^{\sigma_g v} R_g(t) . \tag{2}$$

From equation (28), Section I, we have for the t^{th} vintage that

$$B(1-\alpha)e^{\sigma t}R(t)^{\alpha} - q(t)A(1-\beta)e^{\sigma_g t} R_g(t)^{\beta} = (\sigma - \sigma_g)q(t) .$$

If we plot the distribution of the rents over the vintages, we obtain the following picture (Diagram 1—drawn for $\sigma > \sigma_g$).

Had the two curves $R(v,t)$ and $R_g(v,t)$ been drawn on logarithmic paper, they would have been straight lines, $R(v,t)$ with a slope σ and $R_g(v,t)$ with a slope σ_g.

It is clear that, when $\sigma > \sigma_g$, the rents paid at t to all capital vintages with date of construction v, $v_0(t) < v \leq t$, are higher in the consumer

1. See Section I, equations (26) and (27).

goods sector than in the capital goods sector. For capital vintages with date of construction earlier than $v_0(t)$, the rents are higher in the capital goods sector. We see now that when $\sigma = \sigma_g$, the two distribution curves in the diagram become identical; this means that the rents paid by the two sectors are the same for all vintages of capital. What, then, is the meaning that we could give to σ and σ_g ?

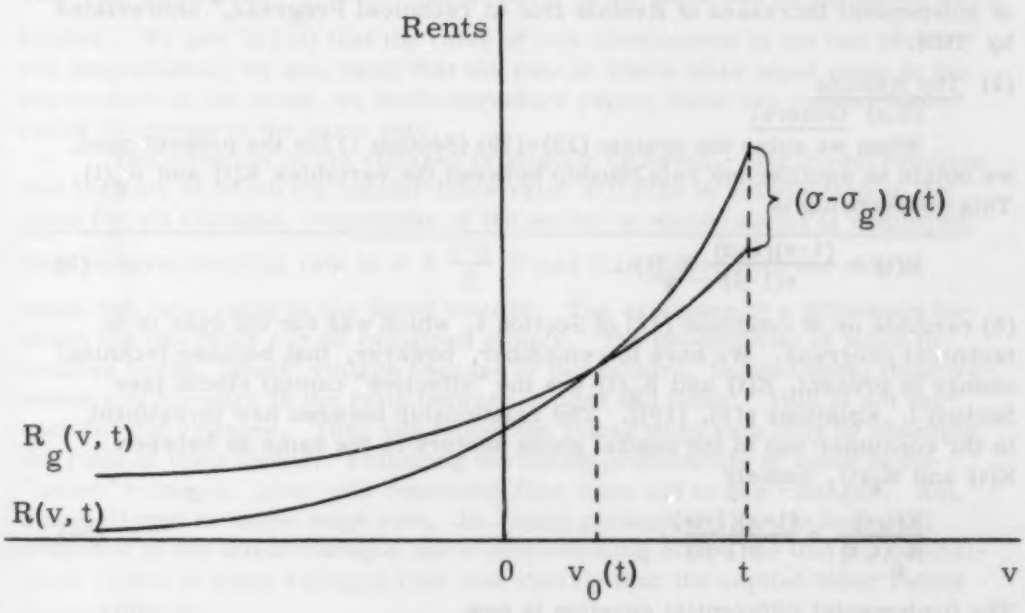

Diagram 1

We can think of the rate at which capital rentals change as depending upon three factors: the change in the wage rate, the change in the price of capital, and the independent change in capital productivity due to technical progress through time.[1] We assert, then, that σ is that component in the rate of change of the rental of new capital employed in the consumer goods sector which is due to increased capital productivity. σ_g is the corresponding component in the rate of increase of rental paid to new capital goods in the capital goods sector. The exact relations are as follows:

$$\frac{\dot{R}(t, t)}{R(t, t)} = \sigma - \frac{\alpha}{1-\alpha} \frac{\dot{W}(t)}{W(t)}, \tag{3}$$

$$\frac{\dot{R}_g(t, t)}{R_g(t, t)} = \sigma_g + \frac{1}{1-\beta} \frac{\dot{q}(t)}{q(t)} - \frac{\beta}{1-\beta} \frac{\dot{W}(t)}{W(t)}. \tag{4}$$

1. Strictly speaking, such a distinction cannot be made since the wage rate, rents, and price of capital all are dependent variables, and they all depend upon the independent parameters of the system. The discussion in the text should be viewed as aimed to strengthen intuition only.

$\frac{\dot{R}(t,t)}{R(t,t)}$ and $\frac{\dot{R}_g(t,t)}{R_g(t,t)}$ are the percentage changes in the rentals paid to a unit of the latest[1] vintage of capital employed in the consumer and capital goods sectors respectively. The proofs for formulae (3) and (4) are given in part (4) of this section.

Because of their special nature we shall call σ and σ_g "The Rates of Independent Increases of Rentals Due to Technical Progress," abbreviated by "RIR."

(2) The Results

(2.a) General

When we solve the system (29)-(39) (Section I) for the present case, we obtain an equilibrium relationship between the variables K(t) and $K_g(t)$. This relationship is

$$K(t) = \frac{(1-s)(1-\alpha)}{s(1-\beta)} K_g(t) . \tag{5}$$

(5) reminds us of condition (43) of Section I, which was for the case of no technical progress. We have to remember, however, that because technical change is present, K(t) and $K_g(t)$ are the "effective" capital stocks (see Section I, equations (18), (19)). The relationship between new investment in the consumer and in the capital goods sectors is the same as between K(t) and $K_g(t)$, namely

$$\frac{K(t,t)}{K_g(t,t)} = \frac{(1-s)(1-\alpha)}{s(1-\beta)} . \tag{5a}$$

The fundamental differential equation is now

$$\frac{\dot{K}_g(t)}{K_g(t)^{1-\beta}} = \frac{As(1-\beta)}{(1-s)(1-\alpha)+s(1-\beta)} \left[\frac{s\beta}{(1-s)\alpha+s\beta}\right]^{\beta} \left[L(0)e^{\left(n+\frac{\sigma}{\beta}\right)t}\right]^{\beta} . \tag{6}$$

After we solve this equation we return to the original system to obtain the time path of all the variables of the system.

On the terminal path, the rate of growth of income is $n + \frac{1-\alpha}{\beta}\sigma$. That is, it is equal to the rate of growth of population plus the expression $\frac{1-\alpha}{\beta}\sigma$, where σ is the common RIR of the two sectors. Obviously, income per capita and consumption per capita are rising on the terminal path at the

---

1. We emphasize the word "latest" since these rates indicate the difference between the rent in vintage t at time t and vintage t+dt at time t+dt. It is not the difference between the rent paid by vintage t at time t and the rent paid by vintage t at time t+dt.

rate $\frac{1-\alpha}{\beta}\sigma$. We then see that the "Iron Law" holds in the present case; the savings ratio does not influence the limiting rate of growth of income.

Since the capital stock is not homogeneous, we cannot speak of "the" capital-labor ratio. Our model shows that on the terminal path the capital-labor ratios in the latest vintages rise at the same rate in the two sectors. This rate is equal to $\frac{1-\beta}{\beta}\sigma$.[1] The equality of the rates can be explained as follows. We saw in (5a) that the rates of new investments in the two sectors are proportional; we also know that the rate at which labor input rises in the two sectors is the same; we would therefore expect these two capital-labor ratios to change at the same rate.

The situation in old vintages is slightly different. Our model predicts that the rate at which the capital-labor ratio will rise in old vintages is the same for all vintages, regardless of the sector in which capital is employed. This uniform limiting rate is $\sigma + \frac{1-\beta}{\beta}\sigma$ and it is higher than the rate at which the ratio rises in the latest vintage. The existence of a difference between the two rates can be explained simply. The productivity of labor increases in the "latest" vintage because of two factors, technological improvements and increase in the ratio between capital and labor. This is not the case for old vintages. They can follow the rising wage rate only by changing the ratio of their inputs. Following the rising productivity of labor in the "latest" vintages, labor will constantly flow from old to new vintages. But, since all pay the same wage rate, the rising productivity due to technical progress in the latest vintages has a compensating effect, so that the capital-labor ratios in these vintages rise less rapidly than the capital-labor ratios in old vintages.

This reveals an interesting fact. The rising wage rate causes a movement of labor from old industries to the new ones; this, in turn, implies a continuous decline in the rent per unit of capital in old vintages. This fact is the basis of Diagram 1 in Part (1) of the present section.

When technical progress is present, the evolution of the wage rate no longer depends solely upon the evolution of the capital labor ratio. It is interesting that when $\alpha > \beta$, the wage rate rises more slowly than the capital-labor ratio in the latest vintage, and when $\beta > \alpha$, it rises more rapidly. We note also that the rates at which the wage rate and the capital-labor ratios rise on the terminal path are independent of the rate of growth of population. This means that the rate of increase of the standard of living of our hypothetical economic community will not depend, in the long run, upon the rate at which this community grows.

The patterns of growth of our economy when it is not on the terminal

1. We shall see in Section III that the rising capital-labor ratio in the latest vintage is due to the existence of technical progress in the capital goods sector only.

path will be summarized with the aid of Diagram 2. Denote by $\bar{Y}(0)$ the initial terminal path level of income. Now let us choose six hypothetical initial levels of income, $Y_1(0)$, $Y_2(0)$, $Y_3(0)$, $Y^1(0)$, $Y^2(0)$, and $Y^3(0)$, such that $Y_3(0) < Y_2(0) < Y_1(0) < \bar{Y}(0) < Y^1(0) < Y^2(0) < Y^3(0)$.[1] That is, $Y_3(0)$ is the most underdeveloped country and $Y^3(0)$ is the most "overdeveloped" country. Now let us denote by $y_i(t)$ and $y^j(t)$ the rates of growth implied by these hypothetical conditions.

In the following diagram, the horizontal line represents the constant rate of growth of income, $n + \frac{1-\alpha}{\beta}\sigma$, on the terminal path. The family of

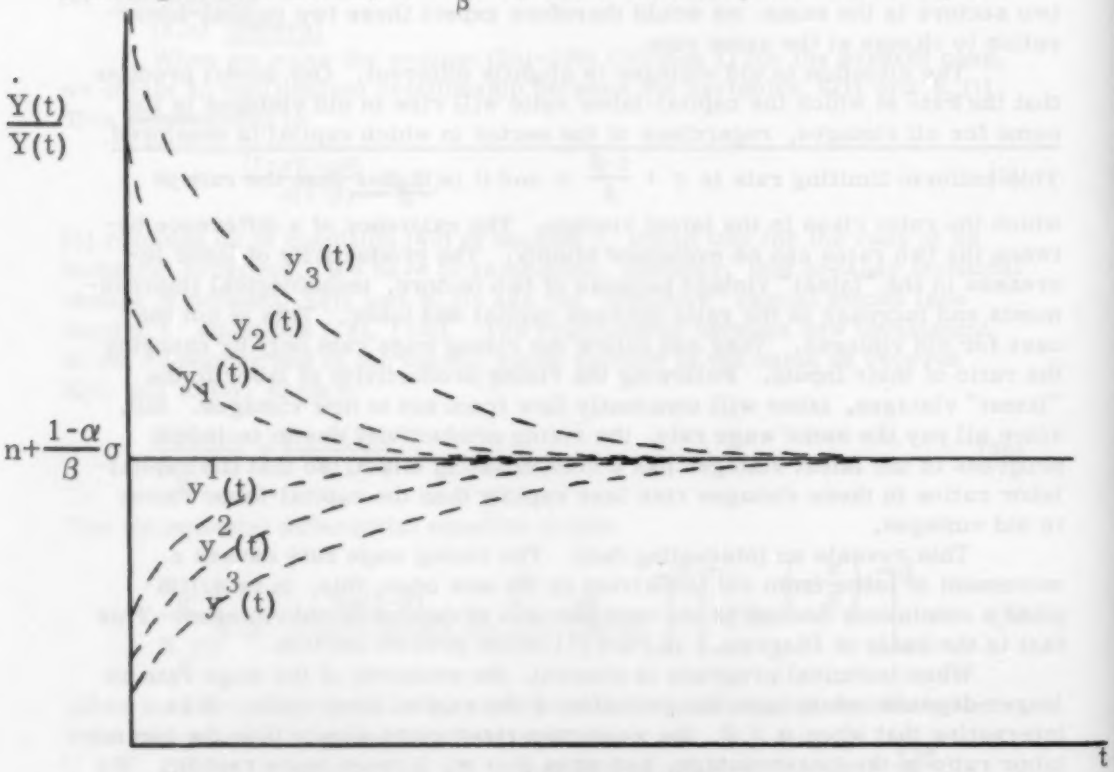

Diagram 2

curves is created by the different initial conditions of the economy. All patterns of growth converge eventually to the terminal path rate of growth.

The diagram shows that the economy with the lowest initial income, namely $Y_3(0)$, will grow at the fastest rate. The economy with the highest initial income, namely $Y^3(0)$, will grow at the slowest rate.[2]

---

1. Here again, $L(0)$ is the same for all these initial conditions. They differ only in their initial capital stocks.

2. For proof, see technical note (4.b) in this section. This proof shows an interesting asymmetry: as initial income tends to zero, its percentage rate of growth tends to infinity, but when the initial income level tends to infinity, its rate of growth does not tend to 0. It tends to αn.

If we examine I(t), income per capita, we find that the competitive equilibrium solution is

$$I(t) = \left[\bar{I}(0)^{\frac{\beta}{1-\alpha}} e^{\sigma t} + \left[I(0)^{\frac{\beta}{1-\alpha}} - \bar{I}(0)^{\frac{\beta}{1-\alpha}} \right] e^{-\beta n t} \right]^{\frac{1-\alpha}{\beta}}.$$

For the underdeveloped country, the initial level I(0) is smaller than the terminal path level of income per capita. It is clear that income per capita will grow on the terminal path at the rate $\frac{1-\alpha}{\beta}\sigma$; before reaching the terminal path, however, I(t) will grow at a rate higher than $\frac{1-\alpha}{\beta}\sigma$. The following diagrams show the level and rate of growth of income per capita.

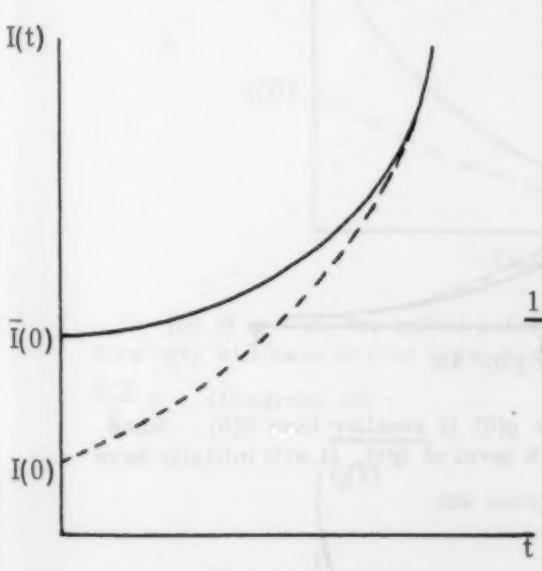

Diagram 3a

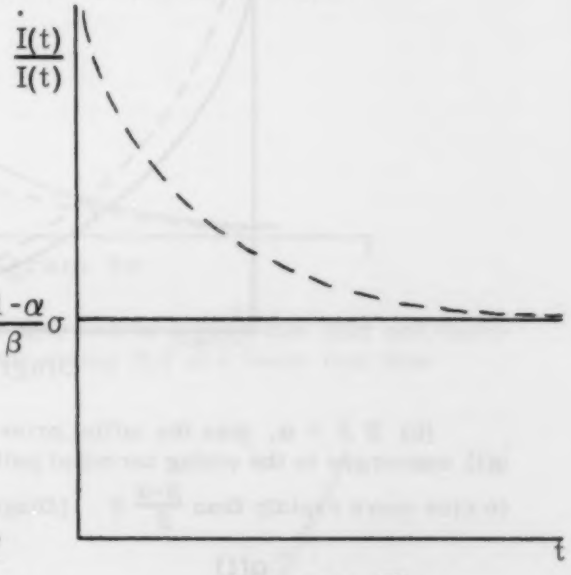

Diagram 3b

The convergence of I(t) to $\bar{I}(t)$ is not always achieved. Whenever I(t) does converge to $\bar{I}(t)$, we say that the model is stable. The stability condition for the case at hand is $\frac{\beta}{1-\alpha} > \frac{\sigma}{\beta n+\sigma}$. Since $\beta n + \sigma \geq \sigma$, a sufficient condition for stability is $\alpha + \beta > 1$.

We can turn now to the valuation side. The presence of technical progress forces the price of new capital goods to change on the terminal path at a rate $\frac{\beta-\alpha}{\beta}\sigma$. We see that the direction of the secular movement of the price depends upon the difference $(\beta-\alpha)$; if $\alpha > \beta$, the price falls, and if $\alpha < \beta$, the price rises. The movement of the price when the economy is not on the terminal path is, however, more complex. Again we start with $\bar{q}(t)$, the terminal path value of the price. We know that $\bar{q}(t) = \bar{q}(0)e^{\frac{\beta-\alpha}{\beta}\sigma t}$.

Two cases are considered.

Case (A) The Underdeveloped Economy

(a) If $\alpha > \beta$, the initial price q(0) will be greater than $\bar{q}(0)$; but since it converges to the falling terminal path value of $\bar{q}(t)$, the actual price q(t) will initially decline more rapidly than $\frac{\beta-\alpha}{\beta}\sigma$. (Diagram 4a)

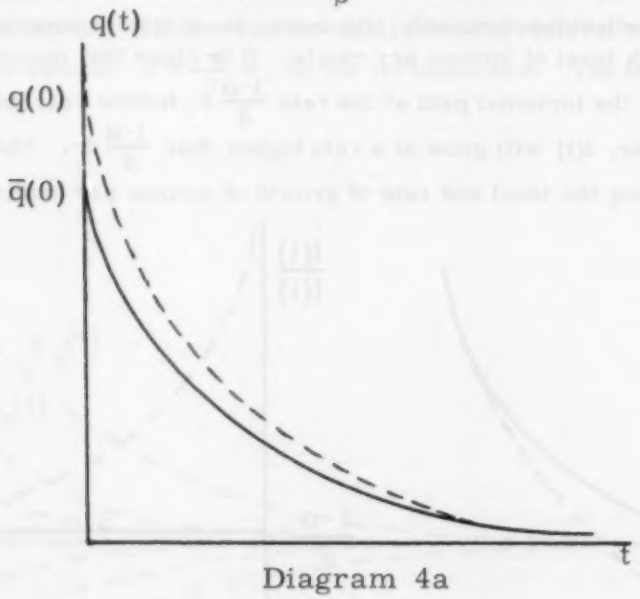

Diagram 4a

(b) If $\beta > \alpha$, then the initial price q(0) is smaller than $\bar{q}(0)$. Since q(t) converges to the rising terminal path level of $\bar{q}(t)$, it will initially have to rise more rapidly than $\frac{\beta-\alpha}{\beta}\sigma$. (Diagram 4b)

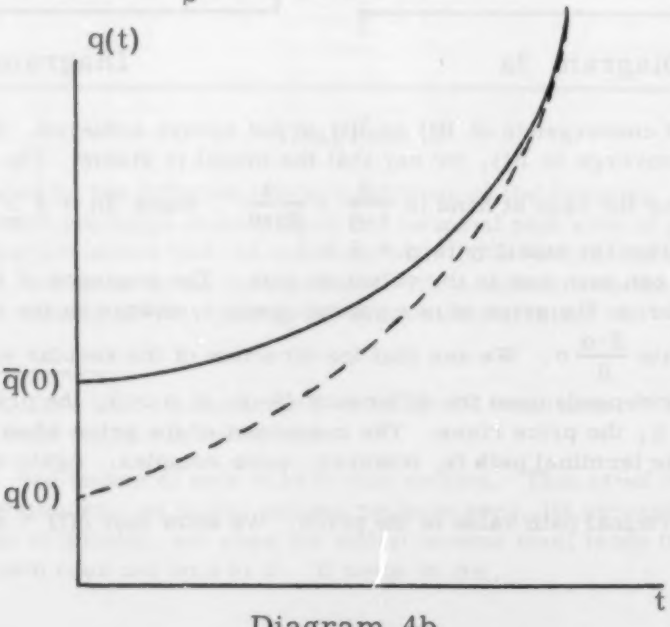

Diagram 4b

Case (B) The Overdeveloped Economy

(a) If $\alpha > \beta$, the initial price q(0) will be smaller than $\bar{q}(0)$ and therefore q(t) will have to decline towards the falling $\bar{q}(t)$ at a lower rate than $\frac{\beta-\alpha}{\beta}\sigma$. (Diagram 4c)

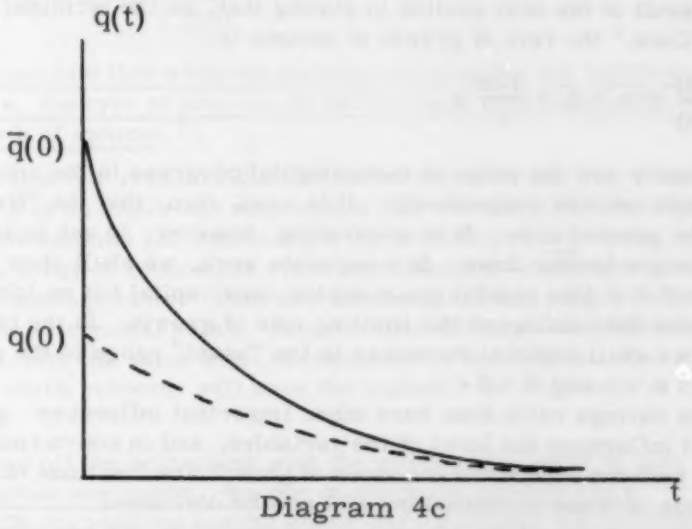

Diagram 4c

(b) If $\alpha < \beta$, the initial price q(0) will be higher than $\bar{q}(0)$ and therefore q(t) will have to rise towards the rising $\bar{q}(t)$ at a lower rate than $\frac{\beta-\alpha}{\beta}\sigma$. (Diagram 4d)

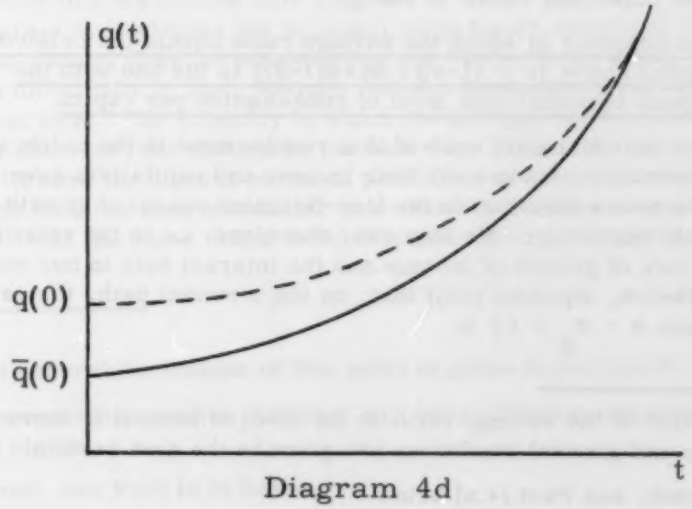

Diagram 4d

Further analysis of the valuation problem (obsolescence, the price of "old" capital goods, and the interest rate) will be presented in Section IV.

(2.b) The Role of the Savings Ratio

The "Iron Law" of growth theory states that the rate of growth of income on the terminal path is independent of the savings ratio. We may anticipate a result of the next section in stating that, on the terminal path of the "General Case," the rate of growth of income is

$$\frac{\dot{\bar{Y}}(t)}{\bar{Y}(t)} = n + \lambda + \frac{1-\alpha}{\beta}\gamma ,$$

where λ and γ are the rates of technological progress in the consumer and capital goods sectors respectively. It is seen, then, that the "Iron Law" holds in the general case. It is interesting, however, to ask under what conditions this law breaks down. In a separate work, we shall show that in the case when $\beta = 0$ (the capital goods sector uses capital but no labor), the savings ratio does influence the limiting rate of growth. In the present work, however, we shall restrict ourselves to the "usual" range of the parameters, namely $0 < \alpha < 1$ and $0 < \beta < 1$.

The savings ratio does have other important influences: on the terminal path it influences the level of the variables, and in non-terminal path situations both the levels and the rates of growth are functions of the savings ratio. Some of these relationships will now be discussed.

Since one aim of the economy is consumption, we are interested in how the savings ratio influences the level of consumption per capita on the terminal path.[1] To answer this question, consider an array of economies which have identical parameters except for their savings ratio. Assume also that all of them are in competitive equilibrium and on their terminal paths. The important result is that[2]

> The economy in which the savings ratio equals the relative capital share $[s = (1-s)(1-\alpha)+s(1-\beta)]$ is the one with the highest terminal path level of consumption per capita.

Note that a special case of this requirement is the celebrated assumption that "workers consume all their income and capitalists save all their rents." We notice also that in the Von-Neumann model of growth this assumption is made implicitly. We may ask, therefore, as to the relationship between the rate of growth of income and the interest rate in our model. We shall see (below, equation (22)) that, on the terminal path, the rate of interest (for the case $\sigma = \sigma_g > 0$) is

---

1. The effect of the savings ratio on the level of income is more complex. Discussion and general conditions are given in the next section.

2. For proof, see Part (4.a) below.

$$\bar{i}(t) = \frac{n}{s}\left[(1-\alpha)(1-s)+(1-\beta)s\right] + \frac{\sigma}{\beta}\left[\frac{(1-\alpha)(1-s)+(1-\beta)s}{s} - \alpha\right]. \qquad (22)$$

Equating the savings ratio to the relative capital share means $s = (1-s)(1-\alpha)+s(1-\beta)$. Using this equality in (22), we have

$$\bar{i}(t) = n + \frac{1-\alpha}{\beta}\sigma = \frac{\dot{\bar{Y}}(t)}{\bar{Y}(t)} .$$

We conclude that when the savings ratio equals the relative capital share, the rate of interest on the terminal path equals the rate of growth of income.[1]

We then see that our result is similar to the result of the Von-Neumann model.

There is an important proposition that we can make about the rate of growth of consumption per capita when the economy is not on the terminal path. Consider an array of economies which are identical in all their parameters except their savings ratio. They are assumed to be in competitive equilibria,[2] but not on their terminal paths. Assume that all these economies start, at time $t = 0$, with the same level of consumption per capita. Our proposition concerning which economy will have the highest rate of growth of consumption per capita is:

among all the economies with the same initial level of consumption per capita, the one in which the savings ratio equals the relative capital share has (uniformly for all time points $t \geq 0$) the highest rate of growth of consumption per capita. In the limit, the rates of growth of all these economies converge to the constant rate $\frac{1-\alpha}{\beta}\sigma$.[3]

We summarize this discussion with diagram 5.

Diagram 5 illustrates the terminal paths (solid lines) and the actual growth paths (dotted lines) of consumption per capita in the array of economies which differ only in their savings ratios. Each path converges to its own terminal path. The economy in which the savings ratio equals the relative capital share is represented by the bold-face lines. We can see that this economy will have the highest terminal path level of consumption per capita, and will approach the path at a faster rate than all the other economies which start from the same initial level of consumption per capita.

---

1. A more general discussion of this point is given in Section IV.

2. This immediately implies that they will have different compositions of the capital stocks.

3. For proof, see Part (4.b) below.

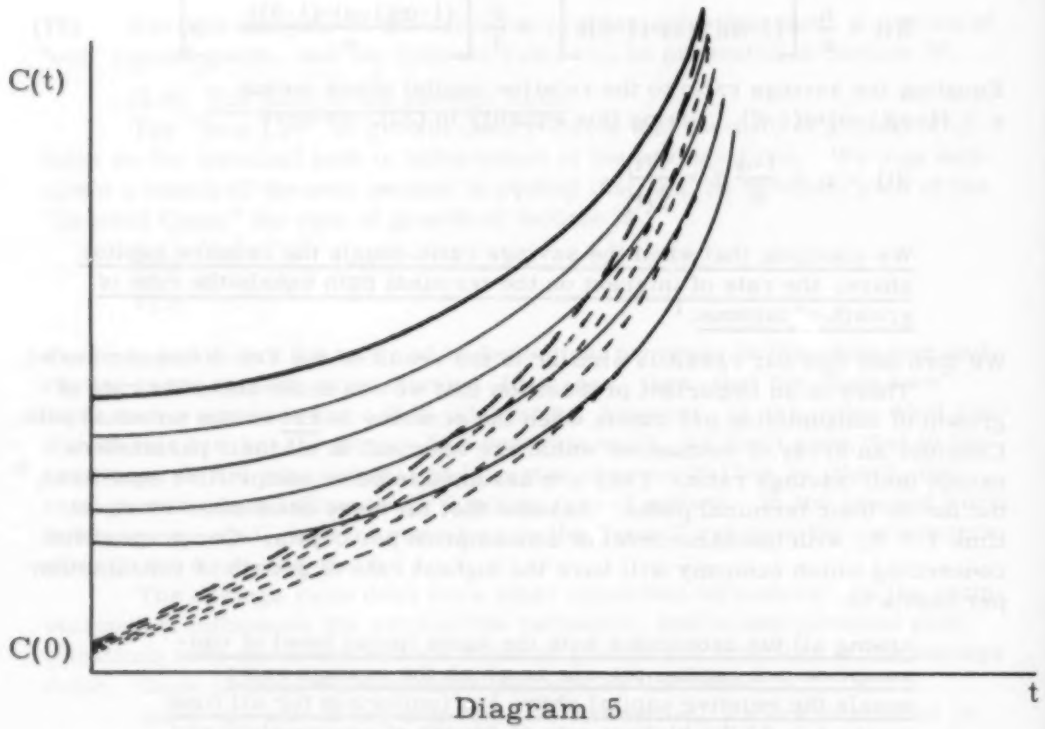

Diagram 5

(3) The Evolution of the Economy through Time in Formulae

Let us again start with income. The solution is

$$Y(t) = \left[\bar{Y}(0)^{\frac{\beta}{1-\alpha}} e^{\frac{\beta}{1-\alpha}(n+\frac{1-\alpha}{\beta}\sigma)t} + \left[Y(0)^{\frac{\beta}{1-\alpha}} - \bar{Y}(0)^{\frac{\beta}{1-\alpha}} \right] e^{\frac{\alpha\beta}{1-\alpha}nt} \right]^{\frac{1-\alpha}{\beta}} \tag{7}$$

and

$$\bar{Y}(0) = \frac{B\alpha}{\alpha(1-s)+\beta s} L(0) \left[\frac{\beta(1-\alpha)}{\alpha(1-\beta)} \right]^{1-\alpha} \left[\frac{As\,(1-\beta)}{s(1-\beta)+(1-s)(1-\alpha)} \frac{1}{n+\frac{\sigma}{\beta}} \right]^{\frac{1-\alpha}{\beta}}, \tag{7a}$$

and clearly

$$\lim_{t \to \infty} \frac{\dot{Y}(t)}{Y(t)} = n + \frac{1-\alpha}{\beta}\sigma \,. \tag{8}$$

We see that the limiting rate of growth of income is equal to the rate of growth of the labor force plus the expression $\frac{1-\alpha}{\beta}\sigma$, where σ is the common RIR of the two sectors. Note that $\sigma = \frac{\lambda}{1-\alpha} = \frac{\gamma}{1-\beta}$; consequently $\frac{1-\alpha}{\beta}\sigma = \frac{\lambda}{\beta}$.

If C(t) is consumption per capita, then

$$C(t) = \left[\bar{C}(0)^{\frac{\beta}{1-\alpha}} e^{\sigma t} + \left[C(0)^{\frac{\beta}{1-\alpha}} - \bar{C}(0)^{\frac{\beta}{1-\alpha}} \right] e^{-\beta n t} \right]^{\frac{1-\alpha}{\beta}} \tag{9}$$

where

$$\bar{C}(0) = \frac{B\alpha(1-s)}{\alpha(1-s)+\beta s} \left[\frac{\beta(1-\alpha)}{\alpha(1-\beta)} \right]^{1-\alpha} \left[\frac{As(1-\beta)}{s(1-\beta)+(1-s)(1-\alpha)} \frac{1}{n+\frac{\sigma}{\beta}} \right]^{\frac{1-\alpha}{\beta}} \tag{9a}$$

and clearly

$$\lim_{t \to \infty} \frac{\dot{C}(t)}{C(t)} = \frac{1-\alpha}{\beta} \sigma . \tag{10}$$

The immediate conclusions we draw are then:

(a) When technological progress is present, the limiting rate of growth of income is higher than the limiting rate of growth of population, so that consumption per capita will eventually be rising at the constant rate $\frac{1-\alpha}{\beta} \sigma$.

(b) The "Iron Law" holds in the present case, i.e., the savings ratio does not affect the limiting rate of growth of income.

We see then, that, as opposed to the case of no technical progress, here there is no upper limit to the amount of consumption per capita that the economy can provide. In Section III, we shall analyze the differences in the contributions of the two rates of technical change to the rising per capita consumption.

We turn now to the capital-labor ratio. When we consider the latest vintage, we shall show that the rate at which capital is invested in this latest vintage increases at a higher rate than the input of labor. This is revealed in the following results. In the consumer goods sector we have

$$r(t, t) = \left[\bar{r}(0, 0)^{\beta} e^{(1-\beta)\sigma t} + \left[r(0, 0)^{\beta} - \bar{r}(0, 0)^{\beta} \right] e^{-\beta(n+\sigma)t} \right]^{\frac{1}{\beta}} , \tag{11}$$

where

$$\bar{r}(0, 0) = \left[\frac{(1-\alpha)\beta}{(1-\beta)\alpha} \frac{As(1-\beta)}{(1-\alpha)(1-s)+s(1-\beta)} \frac{1}{n+\frac{\sigma}{\beta}} \right]^{\frac{1}{\beta}} . \tag{11a}$$

In the capital goods sector we have

$$r_g(t, t) = \left[\bar{r}_g(0, 0)^{\beta} e^{(1-\beta)\sigma t} + \left[r_g(0, 0)^{\beta} - \bar{r}_g(0, 0) \right] e^{-\beta(n+\sigma)t} \right]^{\frac{1}{\beta}} , \tag{12}$$

where

$$\bar{r}_g(0,0) = \left[\frac{As(1-\beta)}{(1-\alpha)(1-s)+s(1-\beta)} \frac{1}{n+\frac{\sigma}{\beta}}\right]^{\frac{1}{\beta}}, \tag{12a}$$

and from (11) and (12) we have that

$$\lim_{t\to\infty} \frac{\dot{r}(t,t)}{r(t,t)} = \lim_{t\to\infty} \frac{\dot{r}_g(t,t)}{r_g(t,t)} = \frac{1-\beta}{\beta}\sigma . \tag{13}$$

We see that the existence of technological change causes the rise in the capital-labor ratio in newly installed vintages. We turn now to the capital-labor ratios in old vintages. From the discussion in Section I (equation (12)) we have that

$$r(v,t) = e^{\sigma(t-v)} r(t,t)$$

and

$$r_g(v,t) = e^{\sigma(t-v)} r_g(t,t) .$$

Consequently, we must have

$$\frac{\dot{r}(v,t)}{r(v,t)} = \sigma + \frac{\dot{r}(t,t)}{r(t,t)} \quad \text{for all } v \tag{14}$$

and

$$\frac{\dot{r}_g(v,t)}{r_g(v,t)} = \sigma + \frac{\dot{r}_g(t,t)}{r_g(t,t)} \quad \text{for all } v \tag{15}$$

and, by (13), we have

$$\lim_{t\to\infty} \frac{\dot{r}(v,t)}{r(v,t)} = \lim_{t\to\infty} \frac{\dot{r}_g(v,t)}{r_g(v,t)} = \sigma + \frac{1-\beta}{\beta}\sigma = \frac{\sigma}{\beta} \text{ for all } v . \tag{16}$$

We have then two conclusions:

(a) The capital-labor ratios in all vintages of both sectors change at the same rate (see equation (16)).

(b) The capital-labor ratio in old vintages increases faster than the capital-labor ratio in the "latest" vintage.

Let us also note that the limiting rate of increase of the capital-labor ratio and of the wage rate (to be seen shortly) does not depend upon n, the rate of growth of the labor force. This, again, is one more aspect of the "Iron Law."

We turn now to the wage rate. The solution in this case is

$$W(t) = \left[\bar{W}(0)^{\frac{\beta}{1-\alpha}} e^{\sigma t} + \left[W(0)^{\frac{\beta}{1-\alpha}} - \bar{W}(0)^{\frac{\beta}{1-\alpha}} \right] e^{-\beta n t} \right]^{\frac{1-\alpha}{\beta}} , \tag{17}$$

$$\bar{W}(0) = \alpha B \left[\frac{(1-\alpha)\beta}{(1-\beta)\alpha} \right]^{1-\alpha} \left[\frac{As(1-\beta)}{(1-\alpha(1-s)+s(1-\beta)} \frac{1}{n+\frac{\sigma}{\beta}} \right]^{\frac{1-\alpha}{\beta}} , \tag{17a}$$

and clearly

$$\lim_{t \to \infty} \frac{\dot{W}(t)}{W(t)} = \frac{1-\alpha}{\beta} \sigma . \tag{18}$$

Note that (18) can be written as

$$\frac{1-\alpha}{\beta} \sigma = \frac{\sigma}{\beta} - \frac{\alpha}{\beta} \sigma = \frac{1-\beta}{\beta} \sigma - \frac{\alpha-\beta}{\beta} \sigma .$$

Hence, by (13), we have

$$\lim_{t \to \infty} \frac{\dot{W}(t)}{W(t)} = \lim_{t \to \infty} \frac{\dot{r}(t,t)}{r(t,t)} - \frac{\alpha-\beta}{\beta} \sigma . \tag{19}$$

This says that when technical progress is present, the limiting rate at which the wage rate rises is equal to the limiting rate at which the capital-labor ratio in the latest vintage rises, minus the expression $\frac{\alpha-\beta}{\beta} \sigma$. The sign of $\frac{\alpha-\beta}{\beta} \sigma$ depends upon the technological differences between the sectors (represented by α and β). If β is smaller than α, the wage rate will rise faster than the capital-labor ratio in the latest vintage.

(3.a) The Price of Capital and the Interest Rate

In the present section we shall state only the results; further analysis is given in Section IV.

The evolution of the price of new capital goods is

$$q(t) = \left[\bar{q}(0)^{\frac{\beta}{\beta-\alpha}} e^{\sigma t} + \left[q(0)^{\frac{\beta}{\beta-\alpha}} - \bar{q}(0)^{\frac{\beta}{\beta-\alpha}} \right] e^{-\beta n t} \right]^{\frac{\beta-\alpha}{\alpha}} \tag{20}$$

where

$$\bar{q}(0) = \frac{B\alpha}{A\beta} \left[\frac{(1-\alpha)\beta}{(1-\beta)\alpha} \right]^{1-\alpha} \left[\frac{As(1-\beta)}{(1-s)(1-\alpha)+s(1-\beta)} \frac{1}{n+\frac{\sigma}{\beta}} \right]^{\frac{\beta-\alpha}{\alpha}} , \tag{20a}$$

and consequently

$$\lim_{t \to \infty} \frac{\dot{q}(t)}{q(t)} = \frac{\beta-\alpha}{\beta}\sigma \ . \tag{21}$$

We see that technical progress will induce a continuous price change through time. This is contrasted with the development program without technical improvements, where the price of capital becomes a constant. The major question is in which direction the price will change. It is clear from (21) that the direction of change depends upon the difference between α and β, the production parameters of the two sectors. If β is smaller than α, there will be a decline in the price. If β is greater than α, the price will rise.[1]

Finally, the limiting rate of interest is

$$\lim_{t \to \infty} i(t) = \frac{n}{s}\left[(1-\alpha)(1-s)+(1-\beta)s\right] + \frac{\sigma}{\beta}\left[\frac{(1-\alpha)(1-s)+(1-\beta)s}{s} - \alpha\right] . \tag{22}$$

Note that the first component is precisely the limiting interest rate when there is no technical progress. The second component arises, obviously, from the presence of such progress.

We note also that in the present case, even if $n = 0$, the interest rate does not become equal to 0. It is possible, however, in the present case, that the interest rate will become negative. The complete analysis is left for Section IV.

(4) Technical Notes and Proofs

(4.a) The Meaning of σ and σ_g

We shall now prove the results stated in equations (3) and (4) and hence the interpretation we have given to σ and σ_g.

We required the wage rate paid by all vintages to be the same; hence,

$$W(t) = \alpha B e^{\lambda v} L(v,t)^{\alpha-1} K(v,t)^{1-\alpha} = \alpha B e^{\lambda t} L(t,t)^{\alpha-1} K(t,t)^{1-\alpha}$$

$$\text{or} \quad \left[e^{\frac{\lambda}{1-\alpha}v} r(v,t)\right]^{1-\alpha} = \left[e^{\frac{\lambda}{1-\alpha}} r(t,t)\right]^{1-\alpha} ,$$

and since $\frac{\lambda}{1-\alpha} = \sigma$, we have

$$e^{\sigma v} r(v,t) = e^{\sigma t} r(t,t) \quad \text{for all } v . \tag{23}$$

This means that, for each t, the product $e^{\sigma v} r(v,t)$ is the same for all v. Indeed, this motivated our transformation $R(t)^{-1} = e^{\sigma t} r(t,t) = e^{\sigma v} r(v,t)$ for all v. Consequently we can write the wage rate

1. The results concerning price behavior when the system is not on the terminal path are obtained by computing the percentage rate of change of $q(t)$ in (20).

$$W(t) = \alpha B R(t)^{\alpha-1} . \tag{24}$$

In the same way, for the capital goods sector we obtain

$$W(t) = \beta A q(t) R_g(t)^{\beta-1} . \tag{25}$$

The rent on new capital in the consumer goods sector is

$$\begin{aligned} R(t,t) &= B(1-\alpha)e^{\lambda t} L(t,t)^{\alpha} K(t,t)^{-\alpha} \\ &= B(1-\alpha)e^{\frac{\lambda}{1-\alpha} t} e^{-\frac{\alpha t}{1-\alpha}} r(t,t)^{-\alpha} \\ &= B(1-\alpha)e^{\sigma t} R(t)^{\alpha} . \end{aligned} \tag{26}$$

From (24) we have

$$\frac{\dot{R}(t)}{R(t)} = \frac{1}{\alpha-1} \frac{\dot{W}(t)}{W(t)} . \tag{27}$$

and from (26) we obtain

$$\frac{\dot{R}(t,t)}{R(t,t)} = \sigma + \alpha \frac{\dot{R}(t)}{R(t)} . \tag{28}$$

Combine (27) and (28) to obtain

$$\frac{\dot{R}(t,t)}{R(t,t)} = \sigma - \frac{\alpha}{1-\alpha} \frac{\dot{W}(t)}{W(t)} . \tag{29}$$

Similarly, the rental in the capital goods sector is

$$R_g(t,t) = q(t)A(1-\beta)e^{\sigma_g t} R_g(t)^{\beta} . \tag{30}$$

From (25) and (30) we have

$$\frac{\beta}{\beta-1} \frac{\dot{W}(t)}{W(t)} = \frac{\beta}{\beta-1} \frac{\dot{q}(t)}{q(t)} + \beta \frac{\dot{R}_g(t)}{R_g(t)} \tag{31}$$

and

$$\frac{\dot{R}_g(t,t)}{R_g(t,t)} = \frac{\dot{q}(t)}{q(t)} + \beta \frac{\dot{R}_g(t)}{R_g(t)} + \sigma_g ; \tag{32}$$

from (31) and (32) we have

$$\frac{\dot{R}_g(t,t)}{R_g(t,t)} = \frac{\dot{q}(t)}{q(t)} - \frac{\beta}{\beta-1} \frac{\dot{q}(t)}{q(t)} + \frac{\beta}{\beta-1} \frac{\dot{W}(t)}{W(t)} .$$

Consequently,

$$\frac{\dot{R}_g(t,t)}{R_g(t,t)} = \frac{1}{1-\beta}\frac{\dot{q}(t)}{q(t)} + \sigma_g - \frac{1}{1-\beta}\frac{\dot{W}(t)}{W(t)} \quad . \tag{33}$$

(29) and (33) establish our contention. It should be clear that σ and σ_g are those components in the rates of change of the rents paid by the latest vintage which are due to technical progress.

(4.b) The Relationship between the Initial Stage of Development and the Initial Rate of Growth

In Part (2), we stated that there is an inverse relationship between the initial rate of growth and the initial stage of development (see Diagram 2). We shall now prove this proposition.

$Y(0)$ is the initial level of income and $\bar{Y}(0)$ is the terminal path initial level of income. Let $y(t)$ be the proportionate rate of growth of income, and let $\mu(0) = \frac{Y(0)}{\bar{Y}(0)}$.

Theorem: The rate of growth of income, $y(t)$, is inversely related to $\mu(0)$. Or: $\frac{dy(t)}{d\mu(0)} < 0$.

Proof: From (7) we compute $y(t)$:

$$y(t) = \frac{1-\alpha}{\beta}\,\frac{\frac{\beta}{1-\alpha}\left(n+\frac{1-\alpha}{\beta}\sigma\right)\bar{Y}(0)^{\frac{\beta}{1-\alpha}}e^{\frac{\beta}{1-\alpha}\left(n+\frac{1-\alpha}{\beta}\sigma\right)t} + \frac{\alpha\beta}{1-\alpha}\left[Y(0)^{\frac{\beta}{1-\alpha}} - \bar{Y}(0)^{\frac{\beta}{1-\alpha}}\right]e^{\frac{\alpha\beta}{1-\alpha}nt}}{\bar{Y}(0)^{\frac{\beta}{1-\alpha}}e^{\frac{\beta}{1-\alpha}\left(n+\frac{1-\alpha}{\beta}\sigma\right)t} + \left[Y(0)^{\frac{\beta}{1-\alpha}} - \bar{Y}(0)^{\frac{\beta}{1-\alpha}}\right]e^{\frac{\alpha\beta}{1-\alpha}nt}} \quad .$$

We then divide through by $Y(0)^{\frac{\beta}{1-\alpha}}e^{\frac{\beta}{1-\alpha}\left(n+\frac{1-\alpha}{\beta}\sigma\right)t}$ using the notation $\mu(0) = \frac{Y(0)}{\bar{Y}(0)}$. We obtain

$$y(t) = \frac{\left(n+\frac{1-\alpha}{\beta}\sigma\right) + \alpha n\left[\mu(0)^{\frac{\beta}{1-\alpha}} - 1\right]e^{-(\beta n+\sigma)t}}{1 + \left[\mu(0)^{\frac{\beta}{1-\alpha}} - 1\right]e^{-(\beta n+\sigma)t}} \quad . \tag{34}$$

Now let $\left[\mu(0)^{\frac{\beta}{1-\alpha}} - 1\right]e^{-(\beta n+\sigma)t} = X(\mu(0))$; (35)

then $$y(t) = \frac{\left(n+\frac{1-\alpha}{\beta}\sigma\right) + \alpha n X(\mu(0))}{1 + X(\mu(0))} \tag{36}$$

hence

$$\frac{dy(t)}{d\mu(0)} = \frac{\partial y(t)}{\partial X} \frac{dX}{d\mu(0)} .$$

From the definition of $X(\mu(0))$ (35) we have

$$\frac{dX}{d\mu(0)} = e^{-(\beta n+\sigma)t} \left(\frac{\beta}{1-\alpha}\right) \mu(0)^{\frac{\beta}{1-\alpha} - 1} > 0$$

and $$\frac{\partial y(t)}{\partial X} = \frac{\alpha n(1+X) - \left(n+\frac{1-\alpha}{\beta}\sigma + \alpha n X\right)}{(1+X)^2} = -\frac{(1-\alpha)n + \frac{1-\alpha}{\beta}\sigma}{(1+X)^2} < 0 ,$$

hence $\frac{dy(t)}{d\mu(0)} < 0$. Q. E. D.

From equation (7) we can obtain the rate of growth of income at $t = 0$. We then have

$$y(0) = \frac{\left(n+\frac{1-\alpha}{\beta}\sigma\right) + \alpha n \left[\mu(0)^{\frac{\beta}{1-\alpha}} - 1\right]}{\mu(0)^{\frac{\beta}{1-\alpha}}} , \qquad (37)$$

and from (37) we see that

$$\lim_{\mu(0) \to 0} y(0) = \infty \qquad (38)$$

and

$$\lim_{\mu(0) \to \infty} y(0) = \alpha n . \qquad (39)$$

Hence, when the ratio $\frac{Y(0)}{\bar{Y}(0)}$ tends to 0, the initial rate of growth tends to ∞, and as the ratio tends to ∞, the initial rate of growth tends to αn . This completes the proof for the validity of the diagram used above.

(4.c) Stability Properties

We stated earlier the conditions for the convergence of $I(t)$, income per capita, to $\bar{I}(t)$, the terminal path level. We shall now prove their validity.

It is clear from the equations we presented that, in the present case, there exists a unique terminal path. To examine the convergence of $\bar{I}(t)$, income per capita, we have

$$I(t) = \left[\bar{I}(0)^{\frac{\beta}{1-\alpha}} e^{\sigma t} + \left[I(0)^{\frac{\beta}{1-\alpha}} - \bar{I}(0)^{\frac{\beta}{1-\alpha}}\right] e^{-\beta n t}\right]^{\frac{1-\alpha}{\beta}} \qquad (40)$$

where

$$\bar{I}(0) = \frac{\beta\alpha}{\alpha(1-s)+\beta s}\left[\frac{\beta(1-\alpha)}{\alpha(1-\beta)}\right]^{1-\alpha}\left[\frac{As(1-\beta)}{(1-s)(1-\alpha)+s(1-\beta)}\ \frac{1}{n+\frac{\sigma}{\beta}}\right]^{\frac{1-\alpha}{\beta}} \cdot \quad (40a)$$

By the definition of stability we say that the system is stable if

$$\lim_{t \to \infty}\left[I(t) - \bar{I}(t)\right] = 0$$

where

$$\bar{I}(t) = \bar{I}(0)e^{\frac{1-\alpha}{\beta}\sigma t} \quad ; \quad (41)$$

hence (40) and (41) yield

$$I(t) = \bar{I}(t)\left[1 + X(t)\right]^{\frac{1-\alpha}{\beta}} \quad (42)$$

where

$$X(t) = \left[\left[\frac{I(0)}{\bar{I}(0)}\right]^{\frac{\beta}{1-\alpha}} - 1\right]e^{-(\beta n+\sigma)t} \quad . \quad (42a)$$

Consequently,

$$I(t) - \bar{I}(t) = \bar{I}(t)\left[(1 + X(t))^{\frac{1-\alpha}{\beta}} - 1\right].$$

Now the limiting value of $\left[I(t) - \bar{I}(t)\right]$ depends upon the rate at which $\bar{I}(t)$ increases, and upon the rate at which $Z(t) = \left[(1+X(t))^{\frac{1-\alpha}{\beta}} - 1\right]$ decreases to 0 .

Now

$$\frac{\dot{Z}(t)}{Z(t)} = \frac{\frac{1-\alpha}{\beta}(1 + X(t))^{\frac{1-\alpha}{\beta} - 1}\frac{dX(t)}{dt}}{(1 + X(t))^{\frac{1-\alpha}{\beta}} - 1} \quad .$$

By applying L'Hospital's rule we find that

$$\lim_{t \to \infty}\frac{\dot{Z}(t)}{Z(t)} = -(\beta n+\sigma) \quad .$$

Since $\frac{\dot{\bar{I}}(t)}{\bar{I}(t)} = \frac{1-\alpha}{\beta}\sigma$, our stability condition is then

$$\beta n + \sigma > \frac{1-\alpha}{\beta}\sigma \quad ,$$

or

$$\frac{\beta}{1-\alpha} > \frac{\sigma}{\beta n + \sigma} \quad , \tag{43}$$

which is the condition stated earlier.

(4.d) The Influence of the Savings Ratio

(d.1) We start by considering the level of consumption per capita. From equations (9)-(10) we have that on the terminal path, $\bar{C}(t)$, the level of consumption per capita, is (for the case $\sigma = \sigma_g \geq 0$)

$$C(t) = \bar{C}(0)e^{\frac{1-\alpha}{\beta}\sigma t}$$

where

$$\bar{C}(0) = \frac{\beta\alpha(1-s)}{\alpha(1-s)+\beta s}\left[\frac{\beta(1-\alpha)}{\alpha(1-\beta)}\right]^{1-\alpha}\left[\frac{As(1-\beta)}{s(1-\beta)+(1-s)(1-\alpha)}\ \frac{1}{n+\frac{\sigma}{\beta}}\right]^{\frac{1-\alpha}{\beta}} .$$

Theorem 1: $\bar{C}(t)$, the level of consumption per capita on the terminal path, attains an absolute interior maximum with respect to s for $0 < s < 1$ at $s^* = s^*(1-\beta)+(1-s^*)(1-\alpha)$.

Proof: Maximization of $\bar{C}(t)$ is equivalent to the maximization of $\bar{C}(0)$ with respect to s. We then take

$$\frac{d \log \bar{C}(0)}{ds} = -\frac{\beta}{(1-s)}\ \frac{1}{\alpha(1-s)+\beta s} + \frac{(1-\alpha)^2}{\beta s[(1-s)(1-\alpha)+s(1-\beta)]}$$

If we let $s^* = (1-s^*)(1-\alpha)+s^*(1-\beta)$

then $1-s^* = \alpha(1-s^*)+\beta s^*$.

Consequently

$$\frac{d \log \bar{C}(0)}{ds} = -\frac{\beta}{(1-s)^2} + \frac{(1-\alpha)^2}{\beta s^2}\ \Bigg|_{s=s^*}$$

$$= \frac{-\beta^2 s^2 + (1-\alpha)^2(1-s)^2}{\beta s^2(1-s)^2}\ \Bigg|_{s=s^*}$$

$$= \frac{((1-\alpha)(1-s)+\beta s)\ ((1-\alpha)(1-s)-\beta s)}{\beta s^2(1-s)^2}\ \Bigg|_{s=s^*} .$$

Clearly $(1-\alpha)(1-s^*) + \beta s^* > 0$; however, $(1-\alpha)(1-s^*) - \beta s^* = 0$ since $0 = (1-s^*)(1-\alpha) + s^*(1-\beta) - s^* = (1-s^*)(1-\alpha) - \beta s^*$.

This proves that when $s^* = (1-s^*)(1-\alpha) + s^*(1-\beta)$, then $\frac{d\bar{C}(0)}{ds} = 0$.

$$\frac{d^2 \log C(0)}{ds} = \frac{\beta[-\alpha(1-s)-\beta s+(1-s)(\beta-\alpha)]}{(1-s)^2[\alpha(1-s)+\beta s]^2}$$

$$-\frac{(1-\alpha)^2}{\beta}\ \frac{(1-s)(1-\alpha)+s(1-\beta)+s(\alpha-\beta)}{s^2[(1-s)(1-\alpha)+s(1-\beta)]^2} \ .$$

Evaluating this expression at $s^* = (1-s^*)(1-\alpha)+s^*(1-\beta)$ yields

$$\frac{d^2 \log \bar{C}(0)}{ds} = -\frac{\beta}{(1-s^*)^3}\ [(1-\beta)+\alpha] \ -\frac{(1-\alpha)^2}{\beta s^{*3}}\ [(1-\beta)+\alpha] \ < 0 \ .$$

Hence $\bar{C}(0)$ attains an interior maximum at $s^* = (1-s^*)(1-\alpha)+s^*(1-\beta)$. To prove that it is the absolute maximum for $0 < s < 1$ consider the expression for $\frac{d \log \bar{C}(0)}{ds}$. Taking a common denominator, the numerator, N(s), is equal to:

$$N(s) = -\beta^2 s + \left[\alpha(1-s)+\beta s\right]\left[\beta^2 s + (1-\alpha)^2(1-s)\right] \ .$$

For $s = 0$, $N(0) > 0$ and for $s = 1$, $N(1) < 0$. But since N(s) is quadratic in s, it can equal zero only at one point for $0 < s < 1$. Q.E.D.

(d.2) Turning now to the rate of growth of an underdeveloped economy, we know from equation (9) that

$$C(t) = \left[\bar{C}(0)^{\frac{\beta}{1-\alpha}} e^{\sigma t} + \left[C(0)^{\frac{\beta}{1-\alpha}} - \bar{C}(0)^{\frac{\beta}{1-\alpha}}\right] e^{-\beta n t}\right]^{\frac{1-\alpha}{\beta}} \ . \qquad (9)$$

Theorem 2: The rate of growth of consumption per capita will attain its absolute interior maximum with respect to s, among all economies with the same initial level of consumption per capita, when $s^* = (1-s^*)(1-\alpha)+s^*(1-\beta)$.

Proof:

$$\frac{d}{dt} \log C(t) = \frac{1-\alpha}{\beta} \frac{\sigma \bar{C}(0)^{\frac{\beta}{1-\alpha}} e^{\sigma t} - \beta n \left[C(0)^{\frac{\beta}{1-\alpha}} - \bar{C}(0)^{\frac{\beta}{1-\alpha}}\right] e^{-\beta n t}}{\bar{C}(0)^{\frac{\beta}{1-\alpha}} e^{\sigma t} + \left[C(0)^{\frac{\beta}{1-\alpha}} - \bar{C}(0)^{\frac{\beta}{1-\alpha}}\right] e^{-\beta n t}}$$

Dividing by $\bar{C}(0)^{\frac{\beta}{1-\alpha}} e^{\sigma t}$ yields

$$\frac{d}{dt}\log C(t) = \frac{1-\alpha}{\beta}\,\frac{\sigma - \beta n\left[\left(\frac{C(0)}{\bar{C}(0)}\right)^{\frac{\beta}{1-\alpha}} - 1\right]e^{-(\beta n+\sigma)t}}{1+\left[\left(\frac{C(0)}{\bar{C}(0)}\right)^{\frac{\beta}{1-\alpha}} - 1\right]e^{-(\beta n+\sigma)t}} \quad . \tag{44}$$

We maximize expression (44) with respect to s.

Let $$X(t) = \left[\left(\frac{C(0)}{\bar{C}(0)}\right)^{\frac{\beta}{1-\alpha}} - 1\right]e^{-(\beta n+\sigma)t} \quad ; \tag{45}$$

hence

$$\frac{d}{dt}\log C(t) = \frac{\frac{1-\alpha}{\beta}\sigma - (1-\alpha)nX(t)}{1+X(t)} \quad ;$$

hence

$$\frac{d}{ds}\left(\frac{d}{dt}\log C(t)\right) = \frac{-(1-\alpha)n\frac{dX(t)}{ds}(1+X(t)) - \frac{dX(t)}{ds}\left[\frac{1-\alpha}{\beta}\sigma - (1-\alpha)nX(t)\right]}{[1+X(t)]^2}$$

$$= \frac{-\frac{dX(t)}{ds}\left[(1-\alpha)n + \frac{1-\alpha}{\beta}\sigma\right]}{[1+X(t)]^2} \quad . \tag{46}$$

Since $(1-\alpha)n + \frac{1-\alpha}{\beta}\sigma$ is positive by assumption, expression (46) will be equal to zero only if $\frac{dX(t)}{ds} = 0$.

From (45),

$$\frac{dX(t)}{ds} = e^{-(\beta n+\sigma)t}\,\frac{\beta}{1-\alpha}\left[\frac{\frac{dC(0)}{ds}\bar{C}(0) - \frac{d\bar{C}(0)}{ds}C(0)}{\bar{C}(0)^2}\right].$$

However, since all economies have the same initial level of consumption per capita, $\frac{dC(0)}{ds} = 0$. Hence

$$\frac{dX(t)}{ds} = \frac{\beta}{1-\alpha}\,e^{-(\beta n+\sigma)t}\left[\frac{C(0)}{\bar{C}(0)^2}\right]\left[-\frac{d\bar{C}(0)}{ds}\right],$$

thus $\frac{dX(t)}{ds}$ will be equal to 0 when $\frac{d\bar{C}(0)}{ds} = 0$. But we know from theorem 1

that $\frac{d\bar{C}(0)}{ds} = 0$ only when $s^* = (1-s^*)(1-\alpha)+s^*(1-\beta)$, hence the rate of growth is also maximized at s^* = capital share. A second derivative test will show the same result as in theorem 1.

III. The General Solution

In Section I we stated that, in principle, the general solution can be obtained for any initial conditions. The great mathematical complexity involved makes it unfeasible, however, to obtain the solutions for non-terminal path initial conditions. In this section, therefore, we present the solutions for the terminal path of the "General Case." In the first part, we state the way in which the results of the general case elaborate on or alter the results of the previous sections. In the second part, the solutions are presented in detail. The proofs are contained in Part (3).

(1) The Results

(1.a) Multiplicity of "Terminal Path" Levels

We learn from the analysis of the general case that the level of the terminal path is not always unique. That is, there are situations in which there are two sets of initial conditions (for $t = 0$) such that if the system starts from either one of them, it will continue to grow at a constant rate. It is the case, however, that the rates of growth are always unique.

To make this statement more precise consider the variable $R_g(t)$. We recall that $R_g(t) = e^{-\sigma_g t} r_g(t, t)^{-1}$. It can be shown that the terminal path solution of $R_g(t)$ is

$$\bar{R}_g(t) = De^{-\frac{\sigma_g}{\beta}} \tag{1}$$

where

$$\bar{R}_g(0) = D = r_g(0,0)^{-1} .$$

The constant D is the terminal path level of the variable $R_g(t)$ at $t = 0$. We shall use this constant in presenting the solutions later in this section. In Part (3.a) below, however, we prove that, in general, the level D is not unique. It follows that the initial terminal path levels of all the variables of the system are not always unique. The exact conditions are stated as follows:

Let $\omega = \sigma - \sigma_g$, then

Case 1: $\omega \geq 0$: In this case there exists a unique level of the terminal path.

Case 2: $\omega < 0$: There are two terminal path levels.[1]

1. The economic meaning of this multiplicity is not clear. It is conjectured here that one of the terminal paths is stable and the other is unstable. There

There is one special case that we shall need in Section IV. If we assume that the savings ratio equals the relative capital share, so that $s = (1-\alpha)(1-s)+(1-\beta)s$, then we have the following two solutions for D^{β} :

$$D_1^{\beta} = \frac{n + \frac{\sigma_g}{\beta}}{A(1-\beta)} \tag{2}$$

and

$$D_2^{\beta} = -\frac{\omega}{A} .^{1} \tag{2a}$$

It is clear that, for the case $\omega > 0$, the level of the terminal path is unique (since we do not accept a negative solution for D). We shall accept D_1^{β} as the level of the terminal path for D^{β}. It is remarkable, however, that D_1^{β} is independent of ω and of any parameter of the consumer goods sector.

(1.b) Asymmetry in the Effects of the Rates of Technical Progress

The rate of growth of income on the terminal path in the general case is

$$\frac{\dot{\bar{Y}}(t)}{\bar{Y}(t)} = n + (1-\alpha)(\omega + \frac{\sigma_g}{\beta}) = n + \lambda + \frac{1-\alpha}{\beta}\gamma .^{2} \tag{3}$$

The following chart helps to see what (3) adds to our knowledge:

Case	Rate of growth of income (terminal path)		
No Technical Change	n	$= n$	
$\sigma = \sigma_g$	$n + \frac{1-\alpha}{\beta}\sigma$	$= n + \frac{\lambda}{\beta}$	$= n + \frac{1-\alpha}{1-\beta} \cdot \frac{\gamma}{\beta}$
General Case	$n + \frac{1-\alpha}{\beta}\sigma_g + (1-\alpha)\omega$	$= n + \lambda + \frac{1-\alpha}{\beta}\gamma$	

We then see that when technical progress is present, the rate of

---

are some indications which allow us to make a conjecture as to which is the stable solution. We cannot, however, offer a conclusive answer to this question since we do not have the solutions for the behavior of the system in non-terminal path situations. The dynamics of a two-sector model in non-terminal path situations is a broad subject still to be investigated.

1. The conjecture (as we mentioned earlier) is that (2a) is the unstable solution. We accept (2) because it will always be a solution regardless of the sign of ω .

2. We remind the reader that λ is the rate of technological progress in the consumer goods sector and γ is the rate of technical progress in the capital goods sector.

growth of income is higher than n. In the case $\sigma = \sigma_g$, we cannot distinguish between the effects of technical progress of the two sectors (note that $n + \frac{1-\alpha}{\beta}\sigma = n + \frac{\lambda}{\beta} = n + \frac{(1-\alpha)\gamma}{(1-\beta)\beta}$). The general case shows us that the rate of growth of income is equal to the rate of growth of the labor force, plus the rate of technical progress in the consumer goods sector, plus the expression $\frac{1-\alpha}{\beta}\gamma$. We see, then, that technical progress in the consumer goods sector contributes λ to the rate of growth of income; technical progress in the capital goods sector contributes $\frac{1-\alpha}{\beta}\gamma$ to the rate of growth of income. Note, however, that this fundamental asymmetry is intimately related to the asymmetric role that the parameters α and β play in the model. We illustrate this with the aid of Diagram 1. Take the rate of growth of income on the terminal path. We have $\frac{\dot{\bar{Y}}}{\bar{Y}} = n + \lambda + \frac{1-\alpha}{\beta}\gamma$. If we plot $\frac{\dot{\bar{Y}}}{\bar{Y}}$ against α and β we have two diagrams:

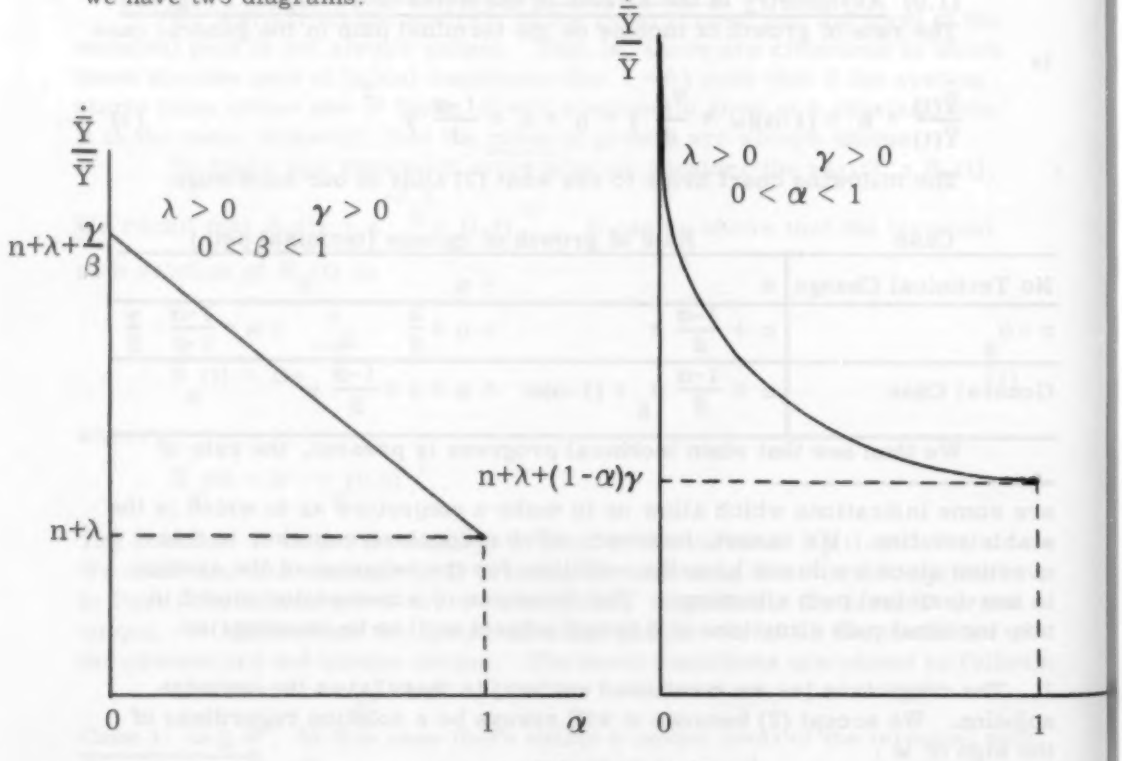

Diagram 1

As α changes from 1 to 0 (for a given β, λ, and γ), the rate of growth of income increases, but it has the upper limit $n + \lambda + \frac{\gamma}{\beta}$. When β tends from 1 to 0, the rate of growth of income becomes explosive.

We should realize that the parameters α and β have a few very important properties. To clarify this point, consider any production

function f(K, L) . We make the following

Definition: For a given production function, the rate at which the marginal productivity of capital diminishes (RDMPC), Q, is

$$Q = \left| \frac{\frac{\partial^2 f}{\partial K^2} K}{\frac{\partial f}{\partial K}} \right| . \qquad (4)$$

Note that $\frac{\partial f}{\partial K}$ is the increase in output due to an increase of one unit of capital. The reduction in the marginal productivity of this factor causes an implicit reduction $\frac{\partial^2 f}{\partial K^2} K$ in capital rents. The smaller Q is, the weaker is the effect of RDMPC on rents. In the extreme case when Q = 0, the marginal productivity of capital is a constant and does not decline. In most models, the RDMPC depends upon the level of inputs or their ratio; we shall see that this is not the case for the Cobb-Douglas production function.

Let Q be the rate of diminishing marginal productivity of capital in the consumer goods sector and let Q_g be the rate in the capital goods sector. It can be shown that

$$Q = \alpha$$
and
$$Q_g = \beta \qquad (5)$$

We then see that both Q and Q_g are independent of the level of inputs, the ratio between capital and labor, and the vintage chosen. They are purely technological measures.

We conclude that the contribution of the rate of technical progress in the capital goods sector (γ) to the rate of growth of income is inversely related to the RDMPC in the capital goods sector. Note that the effect of γ on the rate of growth of income is also inversely related to α, the RDMPC in the consumer goods sector. This, however, has a different effect because the contribution $\frac{1-\alpha}{\beta}\gamma$ is linear in α, but hyperbolic in β.

The asymmetry involved thus has two aspects: the two rates of technological progress have asymmetric contributions to the terminal path rate of growth of income due to different coefficients of λ and γ in $\frac{\dot{\bar{Y}}(t)}{\bar{Y}(t)}$ (equation (3)). The differences in the coefficients reflect the asymmetric role that the parameters α and β play in the model. [1]

1. We shall not attempt to give an intuitive explanation for the basic asymmetry that we are discussing. The following remarks emphasize only the

It might be thought that the asymmetry is due to the choice of the numéraire. This is not the case. The reader may satisfy himself of this by choosing a unit of capital goods of any vintage as a numéraire. Divide the value of income (in terms of consumer goods) by the price of that good (the price can be obtained in Section IV). The asymmetric nature of the results does not change.

(1.c) The Capital-Labor Ratios

We saw in Section II that the capital-labor ratios are not constant on the terminal path when technical progress is present. Let us now compare the "general case" with the case $\sigma = \sigma_g$.

We start with the proportionate rate of change of the capital-labor ratio in the "latest" vintage. It is in general true that the rate is the same for the two sectors. The comparison of the general case with the case $\sigma = \sigma_g$ shows that

in the case $\sigma = \sigma_g$, the rate is $\frac{1-\beta}{\beta}\sigma$, and

for the "general case," the rate is $\frac{1-\beta}{\beta}\sigma_g$.

The true reason for the increase in the capital-labor ratio in the latest vintages is the presence of technical progress in the capital goods sector. If $\sigma_g = 0$, then the ratio in which labor is combined with the latest model of capital goods is the same through time. Only when capital goods become old does the number of workers per machine decline because of a constant flow of workers from old to new vintages.

In Section II ($\sigma = \sigma_g$), the rate at which the capital-labor ratio rises in old vintages is $\frac{\sigma}{\beta}$ for all vintages and for both sectors. In the "general case" the story is different. The percentage rate of change of the capital-labor ratios in "old" vintages is the same within a given sector but different

---

most important elements in the argument.

First, we should observe again the asymmetric role the two goods play in the economy. Consumer goods are consumed and consequently withdrawn from the system. Capital goods are being fed back into the system for further production.

Second, let us observe what happens in the extreme cases. When Q becomes small and tends to 0, the consumer goods sector tends to attract a larger fraction of new capital goods and, eventually, to expand the rate of output more rapidly. This tendency is limited by the RDMPC in the capital goods sector. When Q_g is smaller and tends to 0, the capital goods sector is the one which tends to expand its output. In this case, however, the lack of DMPC in this sector allows this expansion to take place since the expanded output of capital goods is put back into the system for production of more consumer and capital goods. Technological progress in the capital goods sector acts as an accelerator to this process.

for the two sectors. The rate is,

in the consumer goods sector, $\omega + \frac{\sigma_g}{\beta}$, [1] and

in the capital goods sector, $\frac{\sigma_g}{\beta}$.

Hence, if $\omega > 0$, the rate is higher in the consumer goods sector, and if $\omega < 0$, the rate is higher in the capital goods sector. Since $\omega = \sigma - \sigma_g$, the meaning of the above results is that the rate at which workers move toward more modern capital vintages is higher in the sector with the higher RIR. If $\alpha = \beta$, then the rate is higher in the sector with the higher rate of technical progress.

It is important to see that, as the capital-labor ratio rises in a given vintage, the rental paid to a unit of capital falls, and consequently the price of this "old" capital good declines. We shall return to this problem in the next section.

(1.d) The Effect of the Savings Ratio on Income

We shall now consider the influence of the savings ratio on the terminal path level of income, a consideration we ignored when we discussed the effects of the savings ratio in Section II.

Suppose we consider two economies which are identical except for their savings ratio. Both are assumed to be on their terminal path. It is rather interesting that the economy with the higher savings ratio may or may not have a higher level of income on the terminal path. The general conditions are:

(1) If $\omega > 0$ and $\alpha > \beta$, then the economy with the higher savings ratio will always have a higher level of income on the terminal path.

(2) The condition for the economy with the higher savings ratio to have a lower level of income is

$$\frac{\alpha-\beta}{\alpha(1-s)+\beta s} < (1-\alpha)\ \frac{dD}{ds}\ \frac{A(1-\beta)+\omega(1-\beta)D^{-\beta}}{A(1-\beta)D+\omega D^{1-\beta}} \qquad [2,3] \qquad (6)$$

where

$$\frac{dD}{ds} = -\frac{A\,\frac{(1-\alpha)}{s^2}\left[n+\sigma+\frac{1-\beta}{\beta}\sigma_g\right]}{\left[\left(n+\frac{\sigma_g}{\beta}\right)\omega\beta D^{-\beta-1}+A\beta D^{\beta-1}A(1-\beta)\right]} \qquad (6a)$$

1. $\omega = \sigma - \sigma_g$.

2. Note that D is the initial terminal path level of the variable $R_g(t)$. It is stated implicitly by equation (29) in the "technical notes" below.

3. For proof, see the technical notes below.

If we want to understand how this phenomenon is possible, we must remember that income is composed of consumer goods and capital goods. Differences in the savings ratio result in differences in the composition of real output and a different price of capital goods.

One might think that this phenomenon is due to the choice of numéraire (consumer goods). This is not the case. For example, take the case $\sigma = \sigma_g$. If one chooses capital goods in any vintage of either sector as numéraire, then the condition for the level of income (in terms of this new numéraire) to be smaller in the economy with the higher savings ratio is

$$\frac{\alpha - \beta}{\alpha(1-s)+\beta s} < - \frac{(1-\alpha)(1-\beta)}{\beta s[(1-\alpha)(1-s)+s(1-\beta)]} \quad . \tag{7}$$

This condition is independent of the vintage or sector (for the case $\sigma = \sigma_g$).

(2) The Solutions in Formulae

Here we present the actual terminal path solutions for the general case. In the results to be presented, we shall use the symbol D to refer to the initial terminal path value of $R_g(t)$, as discussed in Part (1.a) above.[1]

The terminal path value for income is:

$$\bar{Y}(t) = \bar{Y}(0)e^{n+(1-\alpha)\left(\omega + \frac{\sigma_g}{\beta}\right)} \tag{8}$$

and

$$\bar{Y}(0) = \frac{\alpha B}{\alpha(1-s)+\beta s} L(0) \left[\frac{\left(AD^{\beta} - n - \frac{\sigma_g}{\beta}\right)\beta s}{D\alpha(1-s)\left(n+\frac{1-\beta}{\beta}\sigma_g\right)} \right]^{1-\alpha} \tag{8a}$$

and clearly $\frac{\dot{\bar{Y}}(t)}{\bar{Y}(t)} = n + (1-\alpha)\left(\omega + \frac{\sigma_g}{\beta}\right)$

$$= n + (1-\alpha)\left(\sigma + \frac{1-\beta}{\beta}\sigma_g\right)$$

$$= n + \lambda + \frac{1-\alpha}{\beta}\gamma \, . \tag{9}$$

We conclude that on the terminal path the rate of growth of income is equal to the rate of growth of the labor force (n) plus the rate of technological improvements in the consumer goods sector (λ) plus the expression $\frac{1-\alpha}{\beta}\gamma$, where γ is the rate of technological improvements in the capital goods sector. We note the asymmetry in the contributions of the two rates

1. For implicit solution of D see technical note (3.a), equation (29), below.

of technological progress to the growth of income.

The capital-labor ratios[1] in the two sectors behave as follows: In the consumer goods sector

$$\bar{r}(t,t) = \bar{r}(0,0)e^{\frac{1-\beta}{\beta}\sigma_g t} \tag{10}$$

where

$$\bar{r}(0,0) = \frac{A\beta(1-\alpha)}{\alpha\left[A(1-\beta)D + \omega D^{1-\beta}\right]}, \tag{10a}$$

and in the capital goods sector

$$\bar{r}_g(t,t) = \bar{r}_g(0,0)^{\frac{1-\beta}{\beta}\sigma_g t} \tag{11}$$

and

$$\bar{r}_g(0,0) = \frac{1}{D} . \tag{11a}$$

It is clear, then, that

$$\frac{\dot{\bar{r}}(t,t)}{\bar{r}(t,t)} = \frac{\dot{\bar{r}}_g(t,t)}{\bar{r}_g(t,t)} = \frac{1-\beta}{\beta}\sigma_g = \frac{\gamma}{\beta} . \tag{12}$$

(12) is very significant; it says that r(t,t), the capital-labor ratio in the consumer goods sector, and $r_g(t,t)$, the capital-labor ratio in the capital goods sector, both rise on the terminal path at the same constant rate $\frac{\gamma}{\beta}$, which depends upon the rate of technological improvements in the capital goods sector only. It is inversely related to the parameter β, and as β becomes smaller, the rate at which the capital-labor ratio rises increases.

In a model with technological progress, the wage rate does not depend upon the capital-labor ratio only. The solution for $\bar{W}(t)$ is

$$\bar{W}(t) = \bar{W}(0)e^{(1-\alpha)\left(\sigma + \frac{1-\beta}{\beta}\sigma_g\right)t} \tag{13}$$

and

1. Note that r(t,t) is the capital-labor ratio in the latest vintage of the consumer goods sector. We refer to it as "the capital-labor ratio" since, for any v, the following relation holds: $r(v,t) = e^{\sigma(t-v)}r(t,t)$; hence, when r(t,t) is given, r(v,t) is determined for all v. The relation is similar for the capital goods sector.

$$\bar{W}(0) = \alpha B \left[\frac{\left[AD^{\beta} - n - \frac{\sigma_g}{\beta} \right] \beta s}{\alpha D(1-s) \left[n + \sigma + \frac{1-\beta}{\beta} \sigma_g \right]} \right]^{1-\alpha} \tag{13a}$$

hence

$$\frac{\dot{\bar{W}}(t)}{\bar{W}(t)} = (1-\alpha) \left(\sigma + \frac{1-\beta}{\beta} \sigma_g \right) = \lambda + (1-\alpha) \frac{\gamma}{\beta} \ . \tag{14}$$

We see that, in the limit, the rate at which the wage rate rises depends both upon the rising capital-labor ratio (at the rate $\frac{\gamma}{\beta}$) and upon the independent increase in labor productivity through time.

An important ratio is the composition-of-new-investments ratio, $u(t) = \frac{K(t,t)}{K_g(t,t)}$. In the present solution we have

$$\bar{u}(t) = \frac{AD^{\beta} - \left(n + \frac{\sigma_g}{\beta} \right)}{n + \frac{\sigma_g}{\beta}} \ . \tag{15}$$

We see, then, that this ratio is a constant. We note that when $\sigma = \sigma_g$,

$$\bar{u}(t) = \frac{(1-s)(1-\alpha)}{s(1-\beta)} \ . \tag{15a}$$

If we assume that all profits are saved while all wages are consumed, then

$$\bar{u}(t) = \frac{\beta}{1-\beta} \ . \tag{16}$$

We conclude that the composition of new investments described by the ratio $\bar{u}(t)$ is completely determined by the parameters, and any variations in them will have considerable effect on $\bar{u}(t)$.

The two variables, $K(t)$ and $K_g(t)$, which represent the "effective capital stocks," have the solutions

$$\bar{K}(t) = \bar{K}(0) e^{\left(n + \sigma + \frac{\sigma_g}{\beta} \right) t} \tag{17}$$

where

$$\bar{K}(0) = \frac{A\beta(1-\alpha)(1-s)}{[\alpha(1-s)+\beta s][A(1-\beta)D+\omega D^{1-\beta}]} \ L(0) \tag{17a}$$

and

$$\bar{K}_g(t) = \bar{K}_g(0)e^{\left(n + \frac{\sigma_g}{\beta}\right)t} \tag{18}$$

where

$$\bar{K}_g(0) = \frac{\beta s}{[\alpha(1-s)+\beta s]D} \tag{18a}$$

It follows that

$$\frac{\dot{\bar{K}}(t)}{\bar{K}(t)} = n + \omega + \frac{\sigma_g}{\beta} \quad \text{where} \quad \omega = \sigma - \sigma_g \tag{19}$$

and

$$\frac{\dot{\bar{K}}_g(t)}{\bar{K}_g(t)} = n + \frac{\sigma_g}{\beta} \ . \tag{20}$$

We see that the two "effective capital stocks" do not have to rise at the same rate. Indeed, the sector with the higher RIR will experience a higher rate of growth of its "effective" capital stock.

We now turn to the valuation side. The solution for $\bar{q}(t)$, the price of new capital goods, is

$$\bar{q}(t) = \bar{q}(0)e^{(1-\alpha)\left(\omega + \frac{\sigma_g}{\beta}\right)t + \frac{1-\beta}{\beta}\sigma_g t} \tag{21}$$

where

$$\bar{q}(0) = \frac{B\alpha D^{\alpha-\beta}}{A\beta}\left[\frac{\left(AD^{\beta} - n - \frac{\sigma_g}{\beta}\right)\beta s}{\alpha(1-s)\left(n + \sigma + \frac{1-\beta}{\beta}\sigma_g\right)}\right]^{1-\alpha} . \tag{21a}$$

Clearly

$$\frac{\dot{\bar{q}}(t)}{\bar{q}(t)} = (1-\alpha)\left(\omega + \frac{\sigma_g}{\beta}\right) + \frac{1-\beta}{\beta}\sigma_g = (1-\alpha)\omega + \frac{\beta-\alpha}{\beta}\sigma_g . \tag{22}$$

The prices of "old" capital goods, $a(v,t)$ and $a_g(v,t)$, are obtained from the following equations.[1]

For the consumer goods sector

$$a(v,t) = e^{-\sigma(t-v)}q(t)$$

1. For proof, see Section IV, Part (3).

and for the capital goods sector

$$a_g(v,t) = e^{-\sigma_g(t-v)} q(t) .$$

Finally, the terminal path solution for the interest rate is

$$\bar{i}(t) = (1-\alpha)\omega + \frac{\sigma_g}{\beta}(\beta-\alpha) - \sigma_g + A(1-\beta)D^{\beta} . \tag{23}$$

(3) Technical Notes

In this part, we deal with the multiplicity of terminal path levels and the effect of the savings ratio on the level of income.

(3.a) Multiplicity of "Terminal Path Levels"

Let us start with the variable $R_g(t)$. We recall that $R_g(t) = e^{-\sigma_g t} r_g(t,t)^{-1}$. It can be shown that the terminal path solution for $R_g(t)$ is

$$\bar{R}_g(t) = De^{-\frac{\sigma_g}{\beta}t} \tag{24}$$

where

$$\bar{R}_g(0) = D = r_g(0,0)^{-1} . \tag{24a}$$

Our problem now is to investigate the value of D and its uniqueness. The method that we use for the solution of this problem is as follows.

We have shown in Section I that our original system reduces to the following 4 equations:[1]

$$\dot{K}(t)e^{-\sigma t} + \dot{K}_g(t)e^{-\sigma_g t} = AK_g(t)R_g(t)^{\beta} \tag{25}$$

$$\frac{A\beta(1-\alpha)}{\alpha} e^{\sigma t}R(t) - A(1-\beta)e^{\sigma_g t} R_g(t) = (\sigma - \sigma_g) R_g(t)^{1-\beta} \tag{26}$$

$$R(t)K(t) = \frac{d-1}{d} L(0)e^{nt} \tag{27}$$

$$R_g(t)K_g(t) = \frac{1}{d} L(0)e^{nt} \tag{28}$$

where

$$d = \frac{\alpha(1-s)}{\beta s} + 1 .$$

1. This sub-system is the same as sub-system (40) in Section I; a few changes were made in presenting it.

We use equations (26), (27), and (28) to eliminate K(t), R(t) and $K_g(t)$, and hence we are left with one equation in $R_g(t)$ and $\dot{R}_g(t)$. Now we insert (24) in this equation and put $t = 0$. This leaves us with one equation which involves the parameters of the system and D. This equation is:

$$\frac{A(1-s)(1-\alpha)\left[n+\sigma+\frac{1-\beta}{\beta}\sigma_g\right]}{s\left[A(1-\beta)D+\omega D^{1-\beta}\right]}+\frac{n+\frac{\sigma_g}{\beta}}{D}=\frac{AD^{\beta}}{D} \tag{29}$$

where $\omega = \sigma - \sigma_g$.

Now let

$$d_0 = \frac{A(1-\alpha)(1-s)}{s},$$

$$d_1 = n+\sigma+\frac{1-\beta}{\beta}\sigma_g, \tag{30}$$

and

$$d_2 = A(1-\beta).$$

Equation (29) can then be written as

$$Ad_2D^{2\beta} - (d_0d_1 + (d_1 - \omega)d_2 - A\omega)D^{\beta} - \omega(d_1 - \omega) = 0. \tag{31}$$

Now let $D^{\beta} = y$

and $$E = d_0d_1 + (d_1 - \omega)d_2 - A\omega. \tag{32}$$

Hence (31) is written

$$Ad_2y^2 - Ey - \omega(d_1 - \omega) = 0, \tag{31a}$$

for which the roots are

$$y_1 = \frac{E+\sqrt{E^2+4Ad_2\omega(d_1-\omega)}}{2Ad_2} \tag{33a}$$

and

$$y_2 = \frac{E-\sqrt{E^2+4Ad_2\omega(d_1-\omega)}}{2Ad_2} \tag{33b}$$

Note that $d_0 > 0$, $d_1 > 0$, and $d_2 > 0$, by assumption. Also, $d_1 - \omega = n + \frac{\sigma_g}{\beta} > 0$.

There are three cases to be considered: $\omega > 0$, $\omega = 0$, and $\omega < 0$. Before analyzing these cases, let us observe that $y = \bar{R}_g(0)^\beta$. From equation (25) we have

$$\dot{K}(0) + \dot{K}_g(0) = AK_g(0)R_g(0)^\beta \quad . \tag{25a}$$

Since on any terminal path $\dot{K}(t) > 0$, $\dot{K}_g(t) > 0$ for all t, the left side of (25a) is positive and so is $K_g(0)$. Consequently, any negative solution of $y = \bar{R}_g(0)^\beta$ is not acceptable.

Case 1: $\omega > 0$

In this case, irrespective of the sign of E, there will be one negative and one positive root. This follows since

$|E| < \left| \sqrt{E^2 + 4Ad_2\omega(d_1 - \omega)} \right|$; hence, if $E \geq 0$, then $y_1 > \frac{E}{Ad_2} \geq 0$ and $y_2 < 0$. If $E < 0$, then $y_1 > 0$ and $y_2 < 0$.

We have explained above that $y_2 < 0$ is rejected and the only acceptable solution is $y_1 > 0$.

Case 2: $\omega = 0$

In this case, $E > 0$ and $y_1 = \frac{E}{Ad_2} > 0$ while $y_2 = 0$.

We see that in cases one and two, we have a unique value for the level of the terminal path.

Case 3: $\omega < 0$

From (32) we see that $E > 0$ and $4Ad_2\omega(d_1 - \omega) < 0$, hence $E - \sqrt{E^2 + 4Ad_2\omega(d_1 - \omega)} > 0$. Consequently, (31a) has two positive roots: $y_1 > \frac{E}{Ad_2} > 0$ and $y_2 > 0$.[1] This means that there are two terminal path levels when $\omega < 0$; that is, there exist two sets of initial conditions such that if the system starts from them, it will continue to grow at constant rates.

(3.b) The Influence of the Savings Ratio on the Level of Income

In order to derive the conditions for the sign of $\frac{dY(0)}{ds}$, we observe that it follows from equation (29) that

$$\frac{\left[AD^\beta - n - \frac{\sigma_g}{\beta}\right]\beta s}{D\alpha(1-s)\left(n + \frac{1-\beta}{\beta}\sigma_g\right)} = \frac{A(1-\alpha)\beta\left(n + \sigma + \frac{1-\beta}{\beta}\sigma_g\right)}{\alpha\left(n + \frac{1-\beta}{\beta}\sigma_g\right)}\left[\frac{1}{A(1-\beta)D + \omega D^{1-\beta}}\right] \cdot \tag{34}$$

1. Two complex roots are possible. Investigation of the expression $E^2 + 4Ad_2\omega(d_1 - \omega)$ shows that this possibility is most unlikely.

Using this relation in equation (8a) for the level of income, we obtain

$$\frac{d \log \bar{Y}(0)}{ds} = \frac{\alpha - \beta}{\alpha(1-s)+\beta s} - (1-\alpha)\frac{dD}{ds}\frac{A(1-\beta) + \omega(1-\beta)D^{-\beta}}{A(1-\beta)D + \omega D^{1-\beta}} . \quad (35)$$

Now we need to obtain $\frac{dD}{ds}$. This is obtained from equation (29) again. By differentiating (29) implicitly and solving for $\frac{dD}{ds}$ we obtain

$$\frac{dD}{ds} = - \frac{A\frac{(1-\alpha)}{s^2}\left(n + \sigma + \frac{1-\beta}{\beta}\sigma_g\right)}{\left(n + \frac{\sigma_g}{\beta}\right)\omega\beta D^{-\beta-1} + A\beta D^{\beta-1}A(1-\beta)} . \quad (36)$$

Inserting (36) into (35) we obtain the proportional rate of change of $\bar{Y}(0)$. It is clear that (35) can become negative. Specifying the condition in terms of the variable D, we have that $\frac{d\bar{Y}(0)}{ds} < 0$ if

$$\frac{\alpha - \beta}{\alpha(1-s)+\beta s} > \frac{\frac{dD}{ds}}{D}\frac{A(1-\beta) + \omega(1-\beta)D^{-\beta}}{A(1-\beta) + \omega D^{-\beta}} . \quad (37)$$

Notice that $D = \bar{r}_g(0,0)^{-1}$, where $\bar{r}_g(0,0)$ is the terminal path level of the capital-labor ratio of vintage 0. We would expect an increase in the savings ratio to cause an <u>increase</u> in the capital-labor ratio, and since $\frac{\frac{dD}{ds}}{D} = -\frac{\frac{d\bar{r}_g(0,0)}{ds}}{\bar{r}_g(0,0)}$, we would expect $\frac{\frac{dD}{ds}}{D}$ to be negative.

The condition on the level of income (in the case $\sigma = \sigma_g$) in terms of capital goods is obtained as follows. The level of income in terms of capital goods is obtained in Section II by dividing equation (7a) through by equation (20a).

This yields

$$\bar{Y}_q(0) = A\beta L(0)\left[\frac{A(1-\beta)}{n + \frac{\sigma}{\beta}}\right]^{\frac{1-\beta}{\beta}}\frac{1}{\alpha(1-s)+\beta s}\left[\frac{s}{s(1-\beta)+(1-s)(1-\alpha)}\right]^{\frac{1-\beta}{\beta}} ; \quad (38)$$

hence

$$\frac{d \log \bar{Y}_q(0)}{ds} = \frac{\alpha - \beta}{\alpha(1-s)+\beta s} + \frac{1-\beta}{\beta s}\left(\frac{1-\alpha}{s(1-\beta)+(1-s)(1-\alpha)}\right), \quad (39)$$

hence $\frac{d\bar{Y}_q(0)}{ds} < 0$ if and only if

$$\alpha - \beta < -\frac{(1-\alpha)(1-\beta)}{\beta s}\left(\frac{\alpha(1-s)+\beta s}{s(1-\beta)+(1-s)(1-\alpha)}\right), \tag{40}$$

which is the condition given in the text.

This condition applies to all vintages because all have the same level, except for a multiplicative factor which is independent of s.

IV. Capital Valuation and the Interest Rate

(1) General

In the present section we shall assemble the results concerning the evolution of the price of capital and the interest rate. The behavior of prices when the economy approaches the terminal path has already been analyzed in detail in Section II. Here, we shall concentrate on two other broad subjects:

a) The determinants of the secular behavior of prices of capital goods and the interest rate.

b) The basic relationships between valuation and the conditions of productivity, obsolescence and technical progress, interest rate productivity and obsolescence, and growth of income and the interest rate.

Parts (2) and (3) are devoted to the analysis of prices; Part (4) is devoted to the interest rate and its relationships with other variables of the system.

(2) The Evolution of the Price of New Capital Goods on the Terminal Path

(2.a) General

We have seen that the growth of the labor force and technological improvements are the major forces which maintain the growth of income. When we focus our attention on the evolution of the price of capital and the interest rate, the assumptions concerning entrepreneurial behavior have to be considered. These behavioral elements were introduced in the following way. First we introduced a time function, $i(t)$, that was assumed to be given to all firms and we called it "the interest rate function;" then we required that the "present value" of future rentals be equal to the price of capital wherever capital is used. As time passes, this price changes; however, at every moment of time a price is set such that the capital market is cleared. The evolution of the price of capital thus reflects the changes which take place in the capital market through time.

These changes are caused both by the evolution of the demand for capital goods in the two sectors and by the supply which is provided by the capital goods sector. The evolution of both the demand and the supply are intimately related to the relative profitability of investment in the two sectors and hence to the composition of investment. This is so because changes in the profitability of investment in the two sectors change the demand for capital and cause changes in the composition of new investment. In the long run, the changes in the composition of investment are associated[1]

1. We use the word "associated" because we discuss here the relationships

with another very important outcome: changes occur in both the sectoral composition of the final output and the rate of growth of physical output in both sectors. Hence we see that the examination of the relative productivity of investment in the two sectors, which governs the composition of the demand for capital, leads to important conclusions about the growth of supply of both consumer and new capital goods.

It is a fundamental fact that the relative productivity[1] of investment in the two sectors is determined, on the terminal path, by the technological conditions that govern the production process. It follows that the evolution of the price of capital goods is essentially determined by the technological conditions, which also dictate the relative rates of growth of physical output in the two sectors. This fact is the clue to the understanding of the factors which govern the secular movement of the price of capital. Let us turn now to the analysis of the results.

(2.b) The Results

In Table 1 we present the solutions for six cases.[2] In the first column, we specify the case considered. In the second, $\bar{q}(0)$, the terminal path level at $t = 0$, is presented. Limiting rates of change appear in the third column. The constant D is defined in the last column.

(b.1) An outline of the principal determinants of the secular evolution of the price

As we indicated in Part (2.a) above, the analysis of the secular tendencies of the price of capital is equivalent to the analysis of the evolution of the relative rates of growth of output of the two sectors. This allows us to make the following general proposition:

> The technological conditions that govern the production process determine the rate of growth of physical output in each sector through

---

among endogeneous variables of the system. We clearly are aiming at the discovery of the exogeneous factors which govern the evolution of the prices of capital goods.

1. By "relative productivity," we simply mean the ratio between the marginal productivity of capital invested in the two sectors. This ratio is certainly not independent of the units of measurement; it is only natural that changes in the units will affect the level of the price. It is clear, however, that change in the units will change the price only by a constant. In using the term "technological conditions" we shall, therefore, include the choice of units.

2. We include in the table cases (4) – (6), which are the "extreme" cases: $\beta = 1$, $\beta = o$ and $\alpha = o$. These formulae will not be analyzed. They are presented here to indicate that radical changes take place on the boundary. The comparison of $\alpha = o$ with $\beta = o$ shows that the asymmetry discussed in Section III is also reflected in the valuation side.

TABLE 1

Case	$\bar{q}(0)$	$\frac{\dot{\bar{q}}(t)}{\bar{q}(t)}$	Constant Used
(1) $0 < \beta < 1$ $0 < \alpha < 1$ $\sigma = 0$ $\sigma_g = 0$	$\frac{B(1-\alpha)}{A(1-\beta)}\left[\frac{\alpha(1-\beta)}{\beta(1-\alpha)}\right]^{\alpha}\left[\frac{1}{n}\,\frac{As(1-\beta)}{(1-s)(1-\alpha)+s(1-\beta)}\right]^{\frac{\beta-\alpha}{\beta}}$	0	—
(2) $0 < \beta < 1$ $0 < \alpha < 1$ $\sigma = \sigma_g > 0$	$\frac{B(1-\alpha)}{A(1-\beta)}\left[\frac{\alpha(1-\beta)}{\beta(1-\alpha)}\right]^{\alpha}\left[\frac{As(1-\beta)}{(1-s)(1-\alpha)+s(1-\beta)\,n+\frac{\sigma}{\beta}}\right]^{\frac{\beta-\alpha}{\beta}}$	$\frac{\beta-\alpha}{\beta}\,\sigma$	
(3) $0 < \beta < 1$ $0 < \alpha < 1$ $\sigma > 0$ $\sigma_g > 0$	$\frac{B(1-\alpha)}{A(1-\beta)}\left[\frac{\alpha(1-\beta)}{\beta(1-\alpha)}\right]^{\alpha} D^{\alpha-\beta}\left[\frac{A(1-\beta)}{A(1-\beta)+\omega D^{-\beta}}\right]^{1-\alpha}$	$(1-\alpha)\omega+\frac{\beta-\alpha}{\beta}\,\sigma_g$	Definition of D : $\frac{\frac{A(1-\alpha)(1-s)n}{s\left[A(1-\beta)+\omega D^{-\beta}\right]}}{\sigma} + \frac{A(1-\alpha)(1-s)(\omega+\frac{g}{\beta})}{s\left[A(1-\beta)+\omega D^{-\beta}\right]} = AD^{\beta} - n - \frac{\sigma_g}{\beta}$
(4) $\beta = 1$ $0 < \alpha < 1$ $\sigma > 0$ $\sigma_g > 0$	$\frac{B\alpha}{A}\left[\frac{As}{\alpha(1-s)(\sigma+\gamma+n)}\right]^{1-\alpha}$	$(1-\alpha)\sigma-\alpha\gamma$	—
(5) $\beta = 0$ $0 < \alpha < 1$ $\sigma = 0$ $\sigma_g = 0$	$\frac{Bs}{A(1-s)}\left[\frac{(1-s)(1-\alpha)}{s}\right]^{1-\alpha}\left(\frac{L(0)}{K_g(0)}\right)^{\alpha}$	$a\left[n-\frac{As}{(1-s)(1-\alpha)+s}\right]$	—
(6) $0 < \beta < 1$ $\alpha = 0$ $\sigma = 0$ $\sigma_g = 0$	$\frac{s}{n}\,\frac{B}{(1-s)+(1-\beta)s}$	0	—

their effect on the physical productivity of investments in the two sectors. When these conditions dictate a faster rate of growth of output for the capital goods sector, the price of capital will decline secularly. If the rate of growth of output of the consumer goods sector is higher than the rate of growth of output of capital goods, the price of capital will rise.

In order to analyze the technological conditions which govern the evolution of the price of capital (and also the tendencies specified in the above proposition), we shall analyze the following factors:

(1) Differences between the RDMPC in the two sectors (Q and Q_g).[1]
(2) Differences between the rates of technical progress in the two sectors.
(3) Differences between the two RIR (σ and σ_g).

These effects are not independent since equality in any two of the three implies equality in the third. We shall therefore discuss (2) and (3) together. In addition, we shall examine the effect on price of the rate of growth of the labor force.

(b.2) The impact of differences between the RDMPC

Let us assume here that $\lambda = \gamma = 0$. Then, from the table (case 1), we have

$$\bar{q}(t) = \frac{B(1-\alpha)}{A(1-\beta)} \left[\frac{\alpha(1-\beta)}{\beta(1-\alpha)}\right]^{\alpha} \left[\frac{1}{n} \frac{As(1-\beta)}{(1-s)(1-\alpha)+s(1-\beta)}\right]^{\frac{\beta-\alpha}{\beta}} . \qquad (1)$$

Hence if $A = B$, then equality of Q and $Q_g (\alpha = \beta)$ implies that $\bar{q}(t) = 1$. This is not surprising since when the two sectors are technologically identical, the model becomes essentially a one-sector model.

Let us now take the case where the two sectors experience technical progress but the two RIR are equal ($\sigma = \sigma_g$). In this case (in the table see case 2) we have

$$\bar{q}(t) = \frac{B(1-\alpha)}{A(1-\beta)} \left[\frac{\alpha(1-\beta)}{\beta(1-\alpha)}\right]^{\alpha} \left[\frac{As(1-\beta)}{(1-s)(1-\alpha)+s(1-\beta)} \frac{1}{n+\frac{\sigma}{\beta}}\right]^{\frac{\beta-\alpha}{\beta}} e^{\frac{\beta-\alpha}{\beta}\sigma t} . \qquad (2)$$

It is again clear that if $A = B$ and $Q = Q_g$, then $\bar{q}(t) = 1$. In this case differences between Q and Q_g will completely determine the direction of change of the price of capital goods. If the RDMPC in the capital goods sector is smaller than the corresponding rate in the consumer goods sector ($\beta < \alpha$), the price of new capital goods will fall. When the rate is smaller in the consumer goods sector, the price of new capital goods will rise.

The explanation of this phenomenon is important. When technological

1. The reader is reminded the $Q = \alpha$ and $Q_g = \beta$ are the rates at which the marginal productivity of capital declines in the two sectors. See Section III.

changes are present, the capital stocks in the latest vintages grow faster than the supply of labor, and the result is a constant increase of the capital-labor ratio in these vintages. Since the growth of labor is exogeneously given, the only internal forces which limit growth are the RDMPC in the two sectors. It can be verified that when $\sigma = \sigma_g$, the rents of capital goods are the same in both sectors for all vintages. Consequently, when the RDMPC in the capital goods sector is smaller, this sector absorbs a greater fraction of the available supply of new capital; this causes the output of the capital goods sector to expand at a higher rate than the output of the consumer goods sector, and a secular decline in the price of capital results. When Q is smaller than Q_g, the opposite evolution occurs.

We saw in Section III that Q and Q_g have asymmetric effects on the rate of growth of income. We see here that Q and Q_g have symmetric effects on the direction of change of the price. For the case $\sigma = \sigma_g > 0$, the rate at which $\bar{q}(t)$ changes is $\frac{\beta-\alpha}{\beta}\sigma$. Hence if $Q_g > Q$, the price $\bar{q}(t)$ will rise secularly, and if $Q_g < Q$, the price will fall secularly. There is no symmetry in their quantitative effects on the rate since, from the inspection of $\frac{\beta-\alpha}{\beta}\sigma$, it is clear that the impact of decreasing Q_g is different from the impact of decreasing Q.

(b.3) The impact of technical progress and the RIR

Let us start with the rate of change of $\bar{q}(t)$ on the terminal path of the "general" solution (see case 3 in the table). Here we have

$$\frac{\dot{\bar{q}}(t)}{\bar{q}(t)} = (1-\alpha)\omega + \frac{\beta-\alpha}{\beta}\sigma_g \quad . \tag{3}$$

Suppose now that $Q = Q_g$. Then $\frac{\dot{\bar{q}}(t)}{\bar{q}(t)} = (1-\alpha)\omega$. This means that if the RDMPC are the same in both sectors, the direction of change of the price depends upon the differences of the RIR. If $\sigma > \sigma_g$, then the rental on capital employed in the consumer goods sector rises faster than the rental in the capital goods sector. The demand for capital by the consumer goods sector rises faster than the supply, and the economy consequently experiences a secular increase in the price of capital. The opposite will occur when $\sigma_g > \sigma$.

We have noted that when the RIR are the same, then $\frac{\dot{\bar{q}}(t)}{\bar{q}(t)} = \frac{\beta-\alpha}{\beta}\sigma$. Now instead suppose that $\lambda = \gamma$; then (3) can be rewritten

$$\frac{\dot{\bar{q}}(t)}{\bar{q}(t)} = \lambda - \frac{\alpha}{\beta}\gamma = \frac{\beta-\alpha}{\beta}\lambda = \frac{\beta-\alpha}{\beta}\gamma \quad . \tag{4}$$

Hence when either $\sigma = \sigma_g$ or $\lambda = \gamma$, the direction of change of the price depends upon the difference between the RDMPC in the two sectors.

In the extreme case, when $Q_g = 1$ ($\beta = 1$)[1], we have

1. The model becomes "Austrian" when $\beta = 1$, since in this case the capital goods sector uses only labor.

$$\bar{q}(t) = \frac{B\alpha}{A}\left[\frac{As}{\alpha(1-s)(\sigma+\gamma+n)}\right]^{1-\alpha} e^{(\lambda-\alpha\gamma)t}, \qquad (5)$$

and here the price of new capital goods will rise or fall depending upon the sign of $(\lambda - \alpha\gamma)$.

To summarize, we observe again that there are two factors which can increase the rate of growth of output of a given sector: a reduction of the RDMPC in a sector and an increase in the rate of technical progress in that sector. This means that a smaller RDMPC and a higher rate of technical progress both affect the price of capital in the same direction. Consequently we have to consider three cases:

Case a: $\gamma > \lambda$ and $\beta < \alpha$[1]

In this case, both the RDMPC effect and the technical progress effect are stronger in the capital goods sector; the rate at which the supply of new capital goods increases through time is greater than the rate at which demand increases. Consequently, the price falls secularly.

Case b: $\gamma < \lambda$ and $\beta > \alpha$

The RDMPC effect and the technical progress effect lead to a rate of growth of the demand for capital goods that is greater than the rate of growth of supply; hence, there is a secular increase in the price.

Case c: either $\lambda > \gamma$ but $\beta < \alpha$, or $\lambda < \gamma$ but $\beta > \alpha$

In this case, the secular tendency of the price depends upon the relative strengths of the RDMPC effect and the technical progress effect in the two sectors.

(b.4) The impact of growth of the labor force

It is clear that changes in the rate of growth of the labor force or the savings ratio affect the level of prices on the terminal path, but not their rate of change.

From (1), (2), and (5) we see that the effect of n on $\bar{q}(0)$, the terminal path level of the price, depends also upon the difference between Q and Q_g. When $Q < Q_g$, then $\frac{d\bar{q}(0)}{dn} < 0$, and when $Q > Q_g$, then $\frac{d\bar{q}(0)}{dn} > 0$. This phenomenon should be viewed as the converse of the previous one. There we analyzed the effects of the exogeneous parameters on the rates of growth through their effect on the sectoral composition of investment and hence on the relative supply of capital (output of the capital goods sector); here we observe the impact of increasing the supply of labor. Note that higher RDMPC means lower RDMP of labor! Consequently, it is only natural

1. It can be easily verified that the three cases can be stated equivalently in terms of σ and σ_g: case a: $\sigma_g > \sigma$ and $\beta < \alpha$, case b: $\sigma_g < \sigma$ and $\beta > \alpha$, and case c: $\sigma > \sigma_g$ but $\beta < \alpha$, or $\sigma < \sigma_g$ but $\beta > \alpha$.

that the effect of an increase in the supply of labor on rents and on the price will be the opposite of the effect of an increased supply of capital. The important difference is that changes in the parameters (e.g., rates of technical progress) that lead to changes in the rates of growth of demand for capital and supply of capital change the terminal path rate of change of price; a change in the rate of growth of the labor force, however, changes only the terminal path level of the price.

(3) Obsolescence and the Prices of Old Capital Goods

Let us first recall our assumption that there is no depreciation of capital goods in the model; hence, their productive capability does not change. The changes in their prices are the consequences of changes which take place in the capital market through time. For this reason, we may view our model as a "pure obsolescence" model, where the prices of capital goods change only as a consequence of economic forces and not as the result of physical changes.

We note also that $q(t)$ was defined as the price of "new" capital goods. There is a clear distinction between $a(t,t)$ and $q(t)$. $q(t)$ should be viewed as the price of a new capital good before it is installed; $a(t,t)$ is its price after installation. Obviously, $a(t,t) = q(t)$; however, when we talk about "old" capital goods we refer to those goods which were installed at time $v \leq t$, and $a(v,t)$ is the price, at t, of a capital good which was installed in the consumer goods sector at time v. Similarly, $a_g(v,t)$ is the price at t of a "machine" installed at time v in the capital goods sector.

Definition: The rate of obsolescence, $O(v,t)$, of a capital good of vintage v employed in the consumer goods sector is $O(v,t) = \frac{\dot{a}(v,t)}{a(v,t)}$. The rate of obsolescence in the capital goods sector is $O_g(v,t) = \frac{\dot{a}_g(v,t)}{a_g(v,t)}$.

We define obsolescence to be the change in the price of capital goods after they are installed. The prices $a(v,t)$ and $a_g(v,t)$ are regarded as functions of their installation date v and the time of employment t. $O(v,t)$, for example, can also be defined in the following alternative notation:

$$O(v,t) = \left.\frac{\frac{\partial a(x,y)}{\partial y}}{a(x,y)}\right|_{\substack{x = v \\ y = t}}$$

We shall now proceed to prove some properties of the rates of obsolescence and the prices of old capital goods in our model.

Lemma 1: Under the assumptions of the model (stated in Section I)

$$\frac{a_g(v,t)}{a(v,t)} = e^{\omega(t-v)} \qquad \text{where } \omega = \sigma - \sigma_g .$$

Proof: By the optimality conditions on the use of new capital we have:

$$q(t) = \int_t^{\infty} q(z)A(1-\beta)e^{\sigma_g t} R_g(z)^{\beta} e^{-\int_t^z i(u)du} dz \qquad (6)$$

and

$$q(t) = \int_t^{\infty} B(1-\alpha)e^{\sigma t} R(z)^{\alpha} e^{-\int_t^z i(u)du} dz \ ; \qquad (7)$$

hence

$$\int_t^{\infty} q(z)A(1-\beta)R_g(z)^{\beta} e^{-\int_t^z i(u)du} dz = e^{\omega t}\int_t^{\infty} B(1-\alpha)R(z)^{\alpha} e^{-\int_t^z i(u)du} dz \, . \qquad (8)$$

Now, by definition,

$$a_g(v,t) = \int_t^{\infty} q(z)A(1-\beta)e^{\sigma_g v} R_g(z)^{\beta} e^{-\int_t^z i(u)du} dz \qquad (9)$$

and

$$a(v,t) = \int_t^{\infty} B(1-\alpha)e^{\sigma v} R(z)^{\alpha} e^{-\int_t^z i(u)du} dz \ , \qquad (10)$$

hence

$$a_g(v,t) = e^{\sigma_g v}\int_t^{\infty} q(z)A(1-\beta)R_g(z)^{\beta} e^{-\int_t^z i(u)du} dz \, .$$

By (8),

$$a_g(v,t) = e^{\sigma_g v} e^{\omega t}\int_t^{\infty} B(1-\alpha)R(z)^{\alpha} e^{-\int_t^z i(u)du} dz$$

and, by (10),

$$a_g(v,t) = e^{\sigma_g v} e^{\omega t} e^{-\sigma v} a(v,t),$$

hence

$$\frac{a_g(v,t)}{a(v,t)} = e^{\omega(t-v)} \qquad \text{Q. E. D.}$$

We then see that $\frac{\dot{a}(v,t)}{a(v,t)} - \frac{\dot{a}_g(v,t)}{a_g(v,t)} = \omega$, which is independent of v. Only when $\omega = 0$ are the rates of obsolescence in the two sectors equal.

Now, is it possible that the rates of obsolescence of capital goods of different vintages (but within the same sector) are different?

Theorem 1: Under the assumptions of the model, the rates of obsolescence of capital goods of different vintages within the same sector are the same, regardless of the age of the equipment.

Proof: For the consumer goods sector. Choose some v_0 and v_1 such that $v_0 \neq v_1$. By definition

$$a(v_0,t) = \int_t^\infty B(1-\alpha)e^{\sigma v_0}R(z)^\alpha e^{-\int_t^z i(u)du}\,dz$$

and

$$a(v_1,t) = \int_t^\infty B(1-\alpha)e^{\sigma v_1}R(z)^\alpha e^{-\int_t^z i(u)du}\,dz\ .$$

Hence

$$a(v_0,t)e^{-\sigma v_0} = a(v_1,t)e^{-\sigma v_1}$$

and consequently

$$\frac{\dot{a}(v_0,t)}{a(v_0,t)} = \frac{a(v_1,t)}{a(v_1,t)}\ .$$

The same proof holds for the capital goods sector. Q. E. D.

This theorem tells us that we may write $O(t)$ and $O_g(t)$ instead of $O(v,t)$ and $O_g(v,t)$. It tells us more, however. Suppose we choose $v_0 = t$ and $v_1 = v$; then we have

$$a(v,t) = q(t)e^{-\sigma(t-v)} \tag{11a}$$

and

$$a_g(v,t) = q(t)e^{-\sigma_g(t-v)}\ . \tag{11b}$$

This is because $q(t) = a(t,t) = a_g(t,t)$. Note, however, that

$$\frac{\dot{q}(t)}{q(t)} \neq \frac{\dot{a}_2(t,t)}{a(t,t)}\ .$$ [1]

Indeed, we have that

$$\frac{\dot{q}(t)}{q(t)} - \frac{a_2(t,t)}{a(t,t)} = \sigma \tag{12a}$$

---

1. $a_2(t,t) = \dfrac{\partial a(x,y)}{\partial y}\Big|_{\substack{x=t\\y=t}}$

and

$$\frac{\dot{q}(t)}{q(t)} - \frac{\dot{a}_{g2}(t,t)}{a_g(t,t)} = \sigma_g \quad . \tag{12b}$$

It follows from (12a) and (12b) that

$$O(t) = \frac{\dot{a}(v,t)}{a(v,t)} = \frac{\dot{q}(t)}{q(t)} - \sigma \quad \text{for all } v \tag{13a}$$

and

$$O_g(t) = \frac{\dot{a}_g(v,t)}{a_g(v,t)} = \frac{\dot{q}(t)}{q(t)} - \sigma_g \quad \text{for all } v \quad . \tag{13b}$$

(13a) and (13b) establish the relationships between the rate of change of the price of new capital goods and the rates of obsolescence. It is clear that these rates are not the same in both sectors. From (13a) and (13b) we see that the sector with the higher σ will experience a higher rate of obsolescence.[1] If the two sectors have the same RDMPC, then the sector with the higher rate of technological change will have the higher rate of obsolescence. This result is not surprising, since if the rate of technological improvements, say in the consumer goods sector, is higher, then we would expect capital goods employed in the production of consumer goods to become obsolete more rapidly.

Note that all the results obtained so far in the present part are general and not restricted to the terminal path. Let us now observe some properties of the rates of obsolescence on the terminal path.

One can easily compute from Table 1 above the rates of obsolescence for all the cases considered (use equations (13a) and (13b)). It is seen that the absence of technical changes eliminates the phenomenon of obsolescence. This indicates that, on the terminal path, the process of obsolescence is due only to technical progress.[2] Indeed, if $A = B$, $\alpha = \beta$, and $\sigma = \sigma_g > 0$, i.e., when the units of measurement and the technological conditions of the two sectors are identical (but technical progress is still present in both sectors), and consequently $\bar{q}(t) = 1$, the rates of obsolescence are equal to σ. This is the case that Solow has discussed,[3] in which the price of new capital goods relative to the price of consumer goods is a constant, whereas the

---

1. Usually we think of obsolescence as a process of fall in price; consequently "higher rate of obsolescence" really refers to $-O(t)$ and $-O_g(t)$. This is so since, strictly speaking, a higher σ or σ_g means a smaller $O(t)$ and $O_g(t)$. We shall then speak of "higher rate of obsolescence" in the sense of more rapidly declining values of capital goods.

2. Except for the special case $\beta = 0$, which is not analyzed here.

3. R. M. Solow, "Investment and Technical Progress," op. cit.

price of old capital goods relative to the price of consumer goods (and hence, new capital goods) falls at the rate σ .

We have not examined the signs of the expressions (13a) and (13b). Is obsolescence, as we have defined it, necessarily a process of falling prices of old capital goods? The next theorem clarifies this question for terminal path solutions.

Theorem 2: Under the assumptions of the model, if $0 < \beta < 1$ and $0 < \alpha < 1$, then, on the terminal path, $\bar{O}(t) < 0$.

Proof: The terminal path solution for $\frac{\dot{q}(t)}{q(t)}$ can be obtained from Table 1 above; hence

$$\bar{O}(t) = (1-\alpha)\omega + \frac{\beta-\alpha}{\beta}\sigma_g - \sigma$$

$$= \sigma(1-\alpha-1) + \sigma_g\left[\frac{\beta-\alpha}{\beta} - (1-\alpha)\right]$$

$$= -\alpha\sigma + \sigma_g \frac{\beta\alpha-\alpha}{\beta}$$

$$= -\alpha\sigma - \sigma_g \frac{\alpha}{\beta}(1-\beta) < 0 \quad . \qquad \text{Q. E. D.} \qquad (14)$$

Notice, however, that

$$\bar{O}_g(t) = (1-\alpha)\omega + \frac{\beta-\alpha}{\beta}\sigma_g - \sigma$$

$$= (1-\alpha)\omega + \sigma_g \frac{\beta-\alpha-\beta}{\beta}$$

$$= \lambda - \frac{1-\alpha}{1-\beta}\gamma - \frac{\alpha}{\beta}\sigma_g = \lambda - \gamma\left[\frac{\alpha-2\alpha\beta+\beta}{\beta(1-\beta)}\right] , \qquad (15)$$

and (15) may be positive or negative. Indeed, the condition for (15) to be positive is

$$\lambda > \gamma \frac{\alpha-2\alpha\beta+\beta}{\beta(1-\alpha)} \quad .$$

The above theorem proves that capital goods employed in the consumer goods sector will always lose their value when technical progress is present. It may happen, however, that capital goods employed in the capital goods sector will appreciate in value as time passes! The economic explanation for this fact is as follows: when λ is increased, the rate of growth of demand for capital goods rises. We can always find a λ sufficiently large such that the rate of increase in the demand of the consumer goods sector for new capital goods is so high that even the prices of old machines (which are used to produce the latest model of capital goods) will rise. Note that the value of these goods is measured in terms of consumer goods. It is clear that the relative price of an old machine to the latest is always declining.

In the consumer goods sector, however, there is no such possibility; in terms of consumer goods, the price of old machines producing consumer goods is always declining.

When the economy is not on the terminal path, the solutions for the rates of obsolescence become rather complex and depend upon the initial conditions of the economy. It may be advisable to distinguish between "pure" obsolescence as a secular phenomenon and changes in capital valuation due to the fact that the economy is not on the terminal path.[1] We have avoided making these finer distinctions in this work.

(4) The Interest Rate[2]

In a one-sector model, the interest rate simply equals the marginal productivity of capital. Hence an analysis of the interest rate is equivalent to an analysis of capital productivity. The reader may have noticed that throughout this work we have linked the analysis of the interest rate with the valuation problem. This was not done accidentally: in a two-sector model, the profitability of holding capital depends also upon changes in the price of capital goods relative to consumer goods. Consequently, changes in relative prices will affect the interest rate. Thus, in a two-sector model, the interest rate reflects both the productivity of capital and the changes in the prices of capital goods relative to consumer goods.

(4.a) The Fundamental Relationship

To analyze the interest rate function we need the following

Definition: The rate of rental, $\mu(v,t)$, of a capital good of vintage v employed in the consumer goods sector is the ratio between the value of the rental paid by this instrument and its price. $\mu_g(v,t)$ is defined correspondingly for the capital goods sector.

In symbols:

$$\mu(v,t) = \frac{R(v,t)}{a(v,t)} \qquad (16a)$$

and

1. Indeed, it was suggested that these changes in valuation due to initial conditions should be viewed as capital gains and losses. This is not acceptable since capital gains and losses are usually considered to be the outcomes of uncertain situations. In our model there is no uncertainty.

2. By now it is a well-known fact that the interest rate is not independent of the choice of numéraire. We have chosen a unit of consumer goods to be the numéraire in our model but have not investigated which, if any, of the qualitative results concerning the interest rate change if the numéraire is changed. See Edmond Malinvaud, "Capital Accumulation and Efficient Allocation of Resources," Econometrica, XXI (April 1953), pp. 233-68, and P. Samuelson, Foundations of Economic Analysis (Cambridge: Harvard University Press, 1955), p. 233 (footnote).

$$\mu_g(v,t) = \frac{R_g(v,t)}{a_g(v,t)} \quad . \tag{16b}$$

Theorem 3: Under the assumptions of the model, the rate of rental in each sector is the same for all vintages.

Proof: By definition (see Section I)

$$R(v,t) = B(1-\alpha)e^{\sigma v}R(t)^{\alpha}$$

and

$$a(v,t) = \int_t^{\infty} B(1-\alpha)e^{\sigma v}R(z)^{\alpha}e^{-\int_t^z i(u)du}\,dz \quad .$$

Hence

$$\mu(v,t) = \frac{B(1-\alpha)e^{\sigma v}R(t)^{\alpha}}{\int_t^{\infty} B(1-\alpha)e^{\sigma v}R(z)^{\alpha}e^{-\int_t^z i(u)du}\,dz}$$

$$= \frac{B(1-\alpha)R(t)^{\alpha}}{\int_t^{\infty} B(1-\alpha)R(z)^{\alpha}e^{-\int_t^z i(u)du}\,dz} \quad ,$$

which is independent of v. Similar proof could be constructed for the capital goods sector. Q. E. D.

This theorem shows that we may write $\mu(t)$ and $\mu_g(t)$ instead of $\mu(v,t)$ and $\mu_g(v,t)$. We can now return to the interest rate.

The interest rate was defined implicity (in Section I) by the two equations

$$q(t) = \int_t^{\infty} B(1-\alpha)e^{\sigma t}R(z)^{\alpha}e^{-\int_t^z i(u)du}\,dz \tag{17a}$$

and

$$q(t) = \int_t^{\infty} q(z)A(1-\beta)e^{\sigma_g t}R_g(z)^{\beta}e^{-\int_t^z i(u)du}\,dz \quad . \tag{17b}$$

Since these have to hold for all t, we may observe one implication of the above equations by differentiating them with respect to t. Dividing through by $q(t)$ and transposing terms yields

$$i(t) = \frac{\dot{q}(t)}{q(t)} - \sigma + \frac{B(1-\alpha)e^{\sigma t}R(t)^{\alpha}}{q(t)} \tag{18a}$$

and

$$i(t) = \frac{\dot{q}(t)}{q(t)} - \sigma_g + \frac{q(t)A(1-\beta)e^{\sigma_g t}R_g(t)^{\beta}}{q(t)} \quad . \tag{18b}$$

From (13a) and (13b) we know that

$$O(t) = \frac{\dot{q}(t)}{q(t)} - \sigma \tag{13a}$$

and

$$O_g(t) = \frac{\dot{q}(t)}{q(t)} - \sigma_g \ , \tag{13b}$$

and since $q(t) = a(t,t) = a_g(t,t)$, we have by (16a), (16b), and theorem 3 that

$$\mu(t) = \frac{B(1-\alpha)e^{\sigma t}R(t)^{\alpha}}{q(t)} \tag{19a}$$

and

$$\mu_g(t) = \frac{q(t)A(1-\beta)e^{\sigma_g t}R_g(t)^{\beta}}{q(t)} \quad . \tag{19b}$$

(13a), (13b), (18a), (18b), (19a) and (19b) together yield

$$i(t) = O(t) + \mu(t) \tag{20a}$$

and

$$i(t) = O_g(t) + \mu_g(t) \quad .[1] \tag{20b}$$

This means that the rate of interest is equal to the rate of obsolescence plus the rate of rental in each sector. If, for example, $\sigma > \sigma_g$, then the rate of obsolescence is higher[2] in the consumer goods sector. But, by (20a) and

1. An alternative way of deriving (20a) and (20b) is as follows: At time t, let us invest a unit of consumer goods in buying $\frac{1}{a(v,t)}$ units of capital goods of vintage v employed in the consumer goods sector. Hold these for Δt units of time and sell them at the prevailing price at $(t+\Delta t)$. We shall be earning $\mu(v,t)\Delta t + \frac{a(v,t+\Delta t)}{a(v,t)}$. But a unit of consumer goods today will be worth $1+i(t)\Delta t$ in terms of consumer goods of $t+\Delta t$, if equilibrium is to prevail. We must then have $\mu(v,t)\Delta t + \frac{a(v,t+\Delta t)}{a(v,t)} = 1+i(t)\Delta t$. Dividing through by Δt and taking limits, we obtain $\mu(v,t)+O(v,t) = i(t)$. Using the theorem in the text we can write $\mu(t)+O(t) = i(t)$. For the capital goods sector, the derivation is similar.

2. Recall that we mean that $O(t)$ is more negative than $O_g(t)$.

(20b), we can conclude immediately that the rate of rental in the consumer goods sector must be higher than the corresponding rate in the capital goods sector. This is indeed what we would expect. If the RDMPC in both sectors is the same, then the industry with a higher rate of technological progress will experience a higher rate of rental and at the same time its equipment will become obsolete more rapidly. If $\sigma_g > \sigma$, then the advancing industry is the capital goods industry, which will enjoy a higher rate of rental with a higher rate of obsolescence.

(4.b) The Properties of the Terminal Path Interest Rate

The "Fundamental Relationship" considered in (4.a) above is general. We turn now to examine both the solutions for the terminal path interest rate and their properties in the cases considered.

(b.1) The solutions

The table on the next page contains the limiting values of the interest rate for the same six cases considered in Table 1 above. The two components $O(t)$ and $\mu(t)$ are specified, and the interested reader can easily compute $O_g(t)$ and $\mu_g(t)$ with the aid of equations (13a) and (13b).

Using the solutions of Table 2, we can make a few immediate observations, some of which were stated in earlier sections.

(1) The interest rate always converges to a constant; indeed, the terminal path could have been defined as the state of the economy in which the interest rate is a constant.

(2) The interest rate becomes equal to 0 when the population does not grow and technical progress is not present in either sector.

(3) The higher the savings ratio, the lower is the limiting value of the interest rate. This reminds us of various models in economics which associate a higher supply of savings with a lower interest rate.

(4) The interest rate will, in the limit, be higher the higher is n, the rate of growth of labor. This result would be expected; a higher rate of growth of labor induces the capital-labor ratio to move in favor of higher rates of rentals in both sectors.

(b.2) The relationship between the interest rate and the rate of growth of income

It is clear from Table 2 that there is some relationship between the interest rate and the rate of growth of income.

Definition: The rentals ratio r is the ratio between total rents paid in the economy and total income. [1]

It can be shown that

1. It is obvious from the definition that the rentals ratio and relative capital share are the same. Conceptually, we distinguish between them. See the following remark in the text.

Case	$\lim_{t \to \infty} i(t)$			Remarks
	$\bar{O}(t)$		$\bar{\mu}(t)$	
(1) $0 < \beta < 1$ $0 < \alpha < 1$ $\sigma = 0$ $\sigma_g = 0$			$\frac{n}{s}\left[s(1-\beta)+(1-s)(1-\alpha)\right]$	
(2) $0 < \beta < 1$ $0 < \alpha < 1$ $i = \sigma_g > 0$	$-\frac{\alpha}{\beta}\sigma$	+	$(\beta n+\sigma)\frac{s(1-\beta)+(1-s)(1-\alpha)}{\beta s}$	
(3) $0 < \beta < 1$ $0 < \alpha < 1$ $\sigma > 0$ $\sigma_g > 0$	$(1-\alpha) - \frac{\alpha}{\beta}\sigma$	+	$A(1-\beta)D^{\beta}$	D was defined in Table 1.
(4) $\beta = 1$ $0 < \alpha < 1$ $\sigma > 0$ $\sigma_g > 0$	$-\alpha(\sigma+\gamma)$	+	$(n+\sigma+\gamma)\frac{(1-\alpha)(1-s)}{s}$	
(5) $\beta = 0$ $0 < \alpha < 1$ $\sigma = 0$ $\sigma_g = 0$	$a\left[n - \frac{As}{(1-s)(1-\alpha)+s}\right]$	+	A	
(6) $0 < \beta < 1$ $\alpha = 0$ $\sigma = 0$ $\sigma_g = 0$			$\frac{n}{s}\left[(1-s)+s(1-\beta)\right]$	

$$r = (1-s)(1-\alpha)+s(1-\beta) \ . \qquad (21)$$

Theorem 4: Under the assumptions of the model, if $\beta > 0$ and the savings ratio is equal to the rentals ratio ($s = r$), then, on the terminal path, the interest rate is equal to the rate of growth of income.

Remark: A special case which is applicable here is the celebrated assumption that all profits are saved and all wages are consumed. Here we only require that the savings ratio equals the rentals ratio, which is equal to the relative capital share.

Proof: In Section III, Part (1.a), we presented the roots of D (equations (2) and (2a)). We had for $s = (1-s)(1-\alpha)+s(1-\beta)$

$$D_1^{\beta} = \frac{n + \frac{\sigma_g}{\beta}}{A(1-\beta)} \qquad (22a)$$

and

$$D_2^{\beta} = -\frac{W}{A} \ . \qquad (22b)$$

We have rejected the root D_2^{β}. Now, from Table 2 we have

$$\bar{i}(t) = (1-\alpha)\omega - \frac{\alpha}{\beta}\sigma_g + A(1-\beta)D^{\beta} \ .$$

By (22a) we have

$$\bar{i}(t) = (1-\alpha)\omega - \frac{\alpha}{\beta}\sigma_g + n + \frac{\sigma_g}{\beta} = n + (1-\alpha)\omega + \frac{1-\alpha}{\beta}\sigma_g = \frac{\dot{\bar{Y}}(t)}{\bar{Y}(t)} \ .$$

Notice that if we put $s = s(1-\beta)+(1-s)(1-\alpha)$ in Table 2 for cases (1) and (2) we obtain[1]

$$\bar{i}(t) = \frac{\dot{\bar{Y}}(t)}{\bar{Y}(t)} \ . \qquad \text{Q. E. D.}$$

It is now clear that when the rentals ratio is greater than the savings ratio ($r > s$), the limiting value of the interest rate is higher than the limiting rate of growth of income. When $r < s$, the interest rate is lower than the rate of growth of income on the terminal path.

(b.3) Negative interest rate

We know that the rate of rental is always non-negative; consequently, a negative interest rate can occur only when the rate of obsolescence is higher than the rate of rentals. It is hard to obtain sufficient conditions for a negative interest rate in the general case. For cases (2) and (4), let $n = 0$; then the limiting interest rate is negative if

$$\alpha s > s(1-\beta)+(1-s)(1-\alpha)$$

1. This is also true for cases (4) and (6) that are not analyzed here.

or

$$s > \frac{r}{\alpha} \ ,$$

where we recall that r is the rentals ratio.

If the labor force does grow, the condition is

$$s > \frac{n+\sigma+\gamma}{\sigma+\gamma} \frac{r}{\alpha} \qquad \text{for case (4)}$$

and

$$s > \frac{\beta n+\sigma}{\sigma} \frac{r}{\alpha} \qquad \text{for case (2)} \ .$$

It is clear that a negative interest rate can occur only when technical progress is present. In this case, the configuration of the parameters is such that the rate of obsolescence is greater than the rate of rental, and a negative interest rate becomes necessary to induce holders of capital goods to be willing to continue to hold their assets. Note that if our economy had a third asset, used only as money (cash), then a negative money-interest rate is not possible. Consequently, the price of this asset will have to fall at a constant rate. This means that a constant inflation must occur.

Summary: This study attempts to analyze the effects of inflationary development financing on the economy of an underdeveloped country, in this case Pakistan over the period from 1951 to 1959. The results indicate that deficit financing operations posed no real threat to financial stability before 1955. Prior to this, the financing of public sector development outlays by means of created money definitely contributed to the mobilization of resources for increased capital formation without seriously hampering investment in the private sector. Nonetheless, this mobilization of resources remained small in magnitude. Further, only the presence of large foreign exchange reserves enabled Pakistan to import the capital equipment necessary to the development program.

Following 1955-56, the deficit financing operations were definitely inflationary, and these inflationary pressures, coupled with the accompanying foreign exchange difficulties, frustrated the efforts of the government to obtain additional real resources for capital formation by means of created money. The public sector development program resulted in a curtailment of private investment activity, primarily due to the scarcity of foreign exchange reserves. In general, the analysis indicates that much of the success of inflationary financing of development programs hinges upon the strength of the balance of payments position. In countries in which the import coefficient of investment is high, as is the case with Pakistan, balance of payments difficulties tend, ceteris paribus, to frustrate demand for capital more than consumption demand. This may still occur despite attempts at stringent exchange controls.

This essay is an abridgment of a doctoral dissertation submitted to Yale University in 1961 in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The author is currently an economic advisor to the Saudi Arabian Monetary Agency.

Dissertation Committee: Professors Robert Triffin (Chairman), Lawrence B. Krause, and Mark W. Leiserson.

P. H.

INFLATIONARY FINANCING AND ECONOMIC DEVELOPMENT : THE PAKISTAN EXPERIENCE, 1951-1959

Parvez Hasan

Introduction

1. The Problem

The basic economic problem facing underdeveloped countries in the world today is the poverty of their people. Rapid economic growth which would bring about a material improvement in the living standards of the masses has become a primary objective of the national governments in these countries. It is well recognized that stagnant economies with growing population pressures hold a grave danger not only to the existing national political and social institutions, but also to the stability of world order. However, the task of finding resources for maintaining even a minimum rate of growth in per capita incomes is often stupendous. On the one hand, the taxation base in underdeveloped countries is generally quite narrow, and on the other, the requirements for non-development expenditure such as defense, civil administration, etc., are relatively inflexible. Paradoxically, foreign investments and loans may not be forthcoming until initial large investments in social overhead capital have already been made.

In the face of great difficulties in obtaining resources, as well as the urgent need for the stepping up of the rate of capital formation, deficit finance or money creation, as an instrument for mobilizing real resources for development, offers a temptation which is almost too great to resist. Therefore, it is not surprising that, notwithstanding an awareness of the dangers inherent in inflationary financing of development, the resort to credit creation for meeting budget deficits has been quite widespread in underdeveloped countries in attempts to accelerate the rate of real capital formation. Almost invariably, the result has been the growth of strong inflationary pressures and balance of payments difficulties which, in many cases, have tended to frustrate the objectives of the development programs.

The desperate need for resources in countries struggling to break the vicious circle of poverty and underdevelopment must be appreciated, and attempts of the governments to use all available means to mobilize real resources for development cannot but be considered in a sympathetic light. But the practical question is whether, and under what circumstances, monetary expansion is a suitable instrument for raising these resources. The decision which is required of the policy maker in this regard is always a difficult one. Perhaps greater monetary restraint than is usual would be exercised if there were a clearer understanding of the nature of the inflationary process in underdeveloped economies. Our knowledge of the forces operating on real demand in an inflationary situation in these economies is at present very inadequate indeed. Case studies of countries undertaking deficit financing for economic development are needed and may help to bring about a better understanding of the relationships between inflation and economic growth. The present study is one such attempt.

2. Scope of the Study

It will be the object of this study to analyze the extent and impact of monetary expansion in the Pakistan economy for the years 1951–59, with special reference to the deficit finance operations of the government, which were a major instrument of financing development expenditures in this period. The fiscal year 1951–52 has been selected as the starting point of our study for the reason that it marked the first major jump in government development spending as well as the beginning of a sizable credit creation in the public sector. The fiscal year 1958–59 marks the close of a definite phase. With the coming into power of the revolutionary regime in October 1958, borrowing from the central bank as an instrument for financing public sector outlays was stopped and a concerted effort was made at balancing the budget and controlling inflation. Over the eight years 1951–52 to 1958–59, with which we are concerned, credit creation in the government sector amounted to Rs. 3100–3200 million (nearly equal to the money supply outstanding at the end of March 1951) and helped finance about 55 per cent of the public development outlay, the remaining portion being financed mostly by foreign aid and loans. An over-all appraisal of this extensive use of the monetary instrument in Pakistan in the broad context of the inflationary situation that developed is the main purpose of this study.

For the purposes of this study, deficit financing has been defined as net creation of credit in the government sector. This definition is narrower than the one usually employed in discussions relating to the developed economies, where deficit financing generally refers to the size of the over-all budget deficits. Deficit financing as defined here measures only that part of the budget deficit which has a direct monetary impact. What makes deficit financing a key variable in underdeveloped countries and a major guide to the financial position of the government is the fact that most of the public debt in these countries is held by the banking system, and the possibilities of borrowing from the non-bank sector are often limited. The creation of money by the government is thus very closely related to the size of the over-all budget deficits. Our definition of deficit financing includes government borrowing from the central bank (including declines in government cash balances) as well as from the commercial banks. An implicit assumption here is that borrowings both from the commercial banks and the central bank have net money creating effects in the economy. In other words, the supply of credit by the commercial banks to the government is assumed not to involve a curtailment of their credit to the private sector. This seems to be a particularly realistic assumption in a situation where the government is adding to commercial banks' reserves either by drawing down its cash balances with, or by borrowing from, the central bank. The commercial banks can, in such a situation, extend increasing amounts of credit to the government without curtailing credit availability to the private sector since they are utilizing the reserves which were created by the government in the first place. In another sense, in underdeveloped economies where the central bank often stands ready to extend accommodation to commercial banks

against government securities, the level of private credit is determined largely by the demand for funds.

One of the basic questions to which we shall address ourselves in this enquiry is the effect of continued government deficit financing operations and the resulting inflationary pressures and balance of payments difficulties on the level of real capital formation in Pakistan. It will be assumed that the ultimate target of government policy was to raise over-all investment rather than investment in the public sector alone. The success of the monetary instrument will be considered specifically in the light of the achievement of this objective. The trend in real consumption and investment will be analyzed in terms of the response of the economy to inflationary financing of development outlays. The difficulties faced in squeezing real consumption and in making exchange control effective as an instrument for allocating foreign exchange for investment under conditions of excess demand will then be considered with special reference to the structural characteristics of the Pakistan economy.

3. Main Features of the Pakistan Economy

The income per capita in Pakistan, around $50 per annum, is low even by the standards of less developed countries. It is less than half the level of Egypt, Ceylon, Ghana, and Mexico, and is slightly below that of India. Like most other underdeveloped countries, Pakistan is heavily dependent on agriculture. Agriculture accounted for over 61 per cent of the total national income in 1949–50, provided almost all major exports, and was responsible for the livelihood, directly or indirectly, of 80 per cent of the population. Agriculture, however, despite its importance, is still run on primitive lines. Subsistence farming is the rule, and food crops alone account for over 85 per cent of the total cultivated area; about one-third of the total national income is estimated to originate in the subsistence sector.[1] The balance in food production has been precarious, and in the last four or five years the country has become a net importer of sizable amounts of food products. The low agricultural productivity is reflected not only in the low level of incomes of the farmers, but also in the extremely low yields per acre: rice and wheat yields are among the lowest in the world and barely one-third to one-half of the high levels prevailing in Egypt and Japan.

While agriculture is backward, large-scale manufacturing is of very recent origin. There was little factory industry in Pakistan at the time of the partition (1947); there was practically no processing of even the basic domestic raw materials, i.e., raw jute and raw cotton. Thus, while East Pakistan accounted for over 80 per cent of the entire world supply of jute, the partition boundaries were so drawn that not a single jute mill of the subcontinent came into the province.

Due to the overwhelming dependence on agriculture and the almost

1. Budget of the Government of Pakistan, 1957–58, Economic Survey and Statistics for 1956 (Karachi: Manager, Government of Pakistan Press, 1957), p. 42.

complete absence of large-scale industry, the foreign trade sector has occupied a strategic position in the country's economy, far beyond its proportionate share in the national product. Pakistan has in common with other raw material exporting countries the problem of instability of the export markets, which is heightened by the fact that close to 70 per cent of total foreign exchange earnings are derived from the exports of raw jute and raw cotton. The vulnerability of the economy to changes emanating abroad has been a major argument used in favor of the policy of rapid industrialization which the country has followed since 1950.

The importance of foreign trade is also evident from the heavy dependence of government revenues on foreign trade; the customs duties provide by far the largest single source of income for the central government.

Though essentially Pakistan is a market economy with a large and active private sector, the government exercises a large measure of control—direct as well as indirect—in the strategic sectors. The balance of payments is strictly controlled through an elaborate system of exchange and trade controls. While most exports are free—subject in some cases to export price checks to avoid evasion—imports are subject to licensing, which has been quite restrictive since 1953. The consequence of the present import licensing, applicable to all imports, is that the level of private investment depends directly on the amount of foreign exchange released by the government for capital goods imports. The control of investment is further supplemented by control over capital issues. The government's own outlay on industrial development is carried out through a public corporation (Pakistan Industrial Development Corporation) set up in 1950. In practice, the PIDC has operated in close cooperation with private enterprise.

I. Monetary Expansion in Pakistan: Extent and Impact

In this section, an attempt will be made to provide an over-all view of the monetary and payments developments in Pakistan during 1951–59. In the presentation of the data, use will be made of the system of monetary analysis developed by the O. E. E. C.[1] The reliance on a monetary rather than an income approach in outlining the process of economic expansion and adjustment is necessitated by the extreme inadequacy of the national income accounts (particularly data on consumption) for Pakistan.

The derivation and the significance of the monetary financing equation which forms the basis of the O. E. E. C. model is discussed in sub-section 1. Sub-section 2 outlines, with the help of the financing equation, the nature and the extent of the internal monetary expansion in Pakistan with special reference to the magnitude of money creation by the government and its role in financing development spending in the public sector. The process of inflationary adaptations, in response to the expansionary impulse generated by credit creation in the public sector, is then studied through an analysis of

1. See O. E. E. C. (Organization for European Economic Cooperation), Statistics of Sources and Uses of Finance, 1948–1958 (Paris, 1959).

the liquidity changes, production growth, price increases, and balance of payments deficits. The experience of Pakistan regarding the absorption of monetary expansion through liquidity increases is dealt with at some length in the context of the frequently advanced argument that the rising money-income ratio in a growing economy may provide considerable scope for non-inflationary use of deficit financing.

1. A System of Monetary Analysis

The analytic scheme underlying the O. E. E. C.'s presentation of monetary data is the following equation of the sources and the uses of funds, which implies that total monetary financing (F), the sum of internal credit monetization, external credit monetization, and liquidity changes must be absorbed by production growth, price rises, and the balance of payments deficit:

$$F = Fi + Fe + Fm = \mu_o (y - y_o) p_o + \mu_o (p - p_o) y + D \qquad (1)$$

On the sources side:

Fi = internal credit monetization, which is equal to the increase in banks' domestic assets less the increase in banks' domestic liabilities other than money.

Fe = external credit monetization, which is equal to the increase in non-banks' net external liabilities plus net transfers from abroad.

Fm = relative contraction of monetary liquidity, $(\mu_o - \mu) Y$, where μ_o is the preceding year's money-income ratio, μ is the current monetary ratio, and Y is the national product in current prices.

On the uses side:

$\mu_o (y - y_o) p_o$ measures the increase in money requirements connected with the production growth; $(y - y_o) p_o$ denotes the excess of the current year's national production in real terms (y) over the previous year's production in real terms (y_o) at the preceding year's prices (p_o).

$\mu_o (p - p_o) y$ measures the absorption of funds through price increases; $(p - p_o) y$ denotes the increase in the price level (G. N. P. deflator) in the current year (p) over the price level in the preceding year (p_o), adjusted for the current year's production (y).

D measures the absorption of funds by the excess of imports of goods and services over exports of goods and services. The receipts connected with the external credit monetization are not covered on the credits side and are regarded as a financing item.

The above monetary equation can be easily derived from a simple

version of monetary analysis[1] which equates changes in the money supply (Δ M) with changes in internal banking assets less the increase in banks' domestic liabilities other than money (Fi) minus the decline in net external assets of the banking system (Ne). The decline in the net external assets of the banking system is equal to the over-all external deficit (D) minus external credit monetization (Fe). Thus the sum of Fi and Fe is equal to the sum of the change in money supply (Δ M) and the over-all external deficit (D). It is further evident that the change in money supply (money holdings) can be further apportioned between changes in monetary liquidity and price increases and production growth.[2]

The balance of payments deficit (D), as defined in the above equation, measures the excess of imports of goods and services over exports of goods and services. The current account deficit thus indicated does not have normative significance. The imports financed through net external credit monetization, i.e., through unilateral transfers, foreign investment, and borrowing, need not represent an adjustment to excess demand originating in the economy. A large part of the foreign aid transfers is often autonomous. For purposes of analysis, a narrower definition of the balance of payments deficit is denoted, for instance, by the decline in the external reserves, which are often the main policy variable which can easily be adopted. The exclusion of aid-financed imports and/or external borrowing by the non-bank sector from the uses side of the monetary financing equation would then be matched by a correspondingly narrow definition of external credit monetization on the sources side.

It is evident that the above financing equation has much in common with the equation of exchange. Indeed, one may go so far as to say that the relationship expressed by the former is, in the modified form, necessary for open economies, and is essentially the same relationship as is expressed

---

1. $\Delta M = Fi - Ne$ (2)
   or $Fi = \Delta M + Ne$ (2a)
   Now $Ne = D - Fe$ (3)
   By substituting,
   $Fi + Fe = \Delta M + D$ (4)

2. $\Delta M = \mu Y - \mu_o Y_o$ (5)
   or $\Delta M = \mu_o (Y - Y_o) - (\mu_o - \mu) Y$ (5a)

The first term on the right-hand side in equation (5a) would measure the increase in money requirements if people had retained during the current year the same ratio of money to national product as in the previous year. This increase can be further split into a component $-\mu_o (y - y_o) p_o -$ connected with production growth, and a component $-\mu_o (p - p_o) y -$ connected with price increases.

The second term above, $(\mu_o - \mu) Y$, measures the relative contraction (or expansion) in the public's money holdings and is defined above as Fm.

by the cash balance equation M = kY. It cannot be overemphasized that both equations hold true as identities and that the relationships they depict are merely definitional relationships.

But though the system of monetary analysis outlined above, consisting as it does of ex-post magnitudes, is quite neutral in establishing causal connections, it can be of use in strongly suggesting correlations between important economic variables. Its decided advantage over simpler versions of monetary analysis is that it avoids the pitfalls inherent in excessive emphasis on money supply variations alone. It helps to emphasize that monetary expansion in an open economy need not, if the means of financing balance of payments deficits are available, lead either to an increase in the price level or to a rise in the money supply.[1] Of course, an inevitable element of arbitrariness enters into the concept of money supply, which, in the above equation, is isolated, almost completely, from the other domestic liabilities of the banking system. For purposes of individual countries, however, this does not pose a serious problem because, if the experience warrants, a broader liquidity measurement in lieu of money alone may be substituted in the equation.

2. Application to Pakistan Data, 1951–59

The monetary and balance of payments data for Pakistan for the years 1951–59 are presented in Tables I and II. The tables, which are self-explanatory, are based on the more detailed data contained in Appendix Tables I and II. The necessary qualifications and limitations of the data, particularly of the money national income series, are mentioned in the Appendix, which also discusses the sources and the derivation of the figures. Only two or three main points need to be noted here. The money supply in Pakistan is defined to include currency in circulation, general demand deposits and 'other' deposits (excluding International Monetary Fund Account No. 1 and U.S. Commodity Aid Accounts) with the central bank. We have adhered to this definition for the reason that the ratio of time deposits to demand deposits has shown a reasonable degree of stability, at least in recent years. The net accruals to time deposits held with the commercial banks have, therefore, been regarded as offsets to gross bank lending and are excluded from the figures of internal credit monetization. The external credit monetization has, for the purposes of Table I, been defined as unilateral aid transfers and net external borrowing by the non-bank sector. Table II, however, presents the figures on the basis of a narrower definition of external credit monetization, i.e. excluding unilateral transfers. The exclusion of aid-financed imports from both the sources and uses side of the monetary financing in Table II helps to focus attention on the role of internal credit monetization, which is, in fact, the main autonomous variable with which we are concerned. It may also be mentioned that the figures for national product used

1. The country, in a sense, would be exporting its inflation.

TABLE I

Monetary Developments in Pakistan

(Fiscal Year: April–March. Amounts in millions of rupees. Figures in parentheses in per cent of money supply at the end of the preceding year).

Data for Monetary Analysis	1951–52	1952–53	1953–54	1954–55	1955–56	1956–57	1957–58	1958–59
1. Internal Credit Monetization (Fi)	301 (9.1)	442 (12.4)	297 (9.0)	289 (7.8)	207 (5.3)	625 (13.8)	548 (10.9)	211 (4.0)
2. External Credit Monetization (Fe)	- 2 (0.0)	70 (2.0)	160 (4.8)	264 (7.2)	307 (7.8)	751 (16.6)	714 (14.1)	891 (16.8)
3. Relative Contraction of Monetary Liquidity (Fm)	- 5 (-0.2)	197 (5.5)	-401 (-12.1)	-802 (-21.7)	353 (9.0)	322 (7.0)	- 8 (-0.2)	533 (10.0)
4. Total Monetary Financing (1+2+3) absorbed by:	294 (8.9)	709 (19.9)	56 (1.7)	-249 (-6.7)	867 (22.1)	1698 (37.4)	1254 (24.7)	1635 (30.8)
5. Production Growth $[\mu_o(y-y_o)\,p_o]$	- 3 (-0.1)	145 (4.0)	218 (6.6)	171 (4.6)	-12 (-0.3)	289 (6.3)	4 (0.0)	1 (0.0)
6. Internal Price Increases $[(\mu_o)\,(p-p_o)y]$	256 (7.8)	-182 (-5.1)	-239 (-7.2)	-751 (-20.2)	983 (25.1)	543 (12.0)	250 (4.9)	679 (12.8)
7. Excess of Imports over Exports (D)	35 (1.0)	792 (22.3)	75 (2.2)	315 (8.5)	-188 (-4.8)	924 (20.4)	958 (19.0)	959 (18.1)
8. Suspense Item and Errors and Omissions in Balance of Payments	6 (0.2)	-46 (-1.3)	2 (0.1)	16 (0.4)	84 (2.1)	-58 (-1.3)	42 (0.8)	- 4 (-0.1)

TABLE II

Monetary Developments in Pakistan

(Fiscal Year: April–March. Amounts in millions of rupees. Figures in parentheses in per cent of money supply at the end of the preceding year).

Data for Monetary Analysis	1951–52	1952–53	1953–54	1954–55	1955–56	1956–57	1957–58	1958–59
1. Internal Credit Monetization (Fi)	301 (9.1)	442 (12.4)	297 (9.0)	289 (7.8)	207 (5.3)	625 (13.8)	548 (10.8)	211 (4.0)
2. External Credit Monetization (excluding unilateral aid transfers) (Fe)	- 5 (-0.2)	45 (1.4)	23 (0.8)	208 (5.6)	42 (1.0)	39 (0.8)	- 6 (-0.1)	57 (1.1)
3. Relative Contraction of Monetary Liquidity (Fm)	- 5 (-0.2)	197 (5.5)	-401 (-12.1)	-802 (-21.7)	353 (9.0)	322 (7.0)	- 8 (-0.2)	533 (10.0)
4. Total Monetary Financing absorbed by:	291 (8.8)	684 (19.3)	-79 (-2.3)	-305 (-8.3)	602 (15.3)	986 (21.6)	534 (10.5)	801 (15.1)
5. Production Growth $[\mu_o(y-y_o)p_o]$	- 3 (-0.1)	145 (4.0)	218 (6.6)	171 (4.6)	-12 (-0.3)	289 (6.3)	4 (0.0)	1 (0.0)
6. Internal Price Increases $[\mu_o(p-p_o)y]$	256 (7.8)	-182 (-5.1)	-239 (-7.2)	-751 (-20.2)	983 (23.1)	543 (12.0)	250 (4.9)	679 (12.8)
7. Excess of Imports over Exports (Excluding aid-financed imports)	32 (0.9)	767 (21.7)	-60 (-1.8)	259 (6.9)	-453 (-11.6)	212 (4.6)	238 (4.8)	125 (2.4)
8. Suspense Item and Errors and Ommissions in Balance of Payments	6 (0.2)	-46 (-1.3)	2 (0.1)	16 (0.4)	84 (2.1)	-58 (-1.3)	42 (0.8)	- 4 (-0.1)

in the estimates are those of national income at factor cost. No allowance has been made for the terms of trade effect, which has been a persistent factor making for a decline in real income in Pakistan in recent years. The production increase is thus somewhat exaggerated, but this qualification in no way invalidates the main conclusions.

Two Distinct Periods

From the point of view of monetary developments in Pakistan, the period from 1951–59 can be seen as falling into two distinct phases. While internal credit monetization was at a high level thoughout, the behavior of the other important variables, i.e., the liquidity ratio, prices, and the balance of payments, showed a marked divergence between 1951–55 and 1955–59. For instance, during 1951–55, the increase in liquidity (a decrease in income velocity) offset more than half the current credit monetization. In sharp contrast, in the next four-year period, the opposite movement in liquidity provided a strong supplement to credit monetization, with the result that total monetary financing increased markedly. On the absorption side, external deficits alone, during 1951–55, exceeded the total monetary financing; prices actually declined rather sharply and, for the period as a whole, the decline in money requirements on this account was more than enough to meet the needs of production growth. With the exception of 1955–56, heavy balance of payments deficits continued to absorb large amounts of monetary financing during the period 1955–59; prices rose significantly, accounting for 45 per cent of the total monetary financing, with absorption by production growth being only 5 per cent. If we adopt a narrow definition of the balance of payments deficit by excluding the imports financed by aid transfers, we find that the current account balance of payments, which, even after excluding unilateral transfers, had absorbed almost the entire monetary financing during 1951–55, showed a rough balance during 1955–59. The expansionary impact of internal credit monetization and contraction in liquidity was thus not accommodated by a balance of payments deficit in the latter period, and excess demand took the form almost entirely of price increases.

Credit Creation and Budget Deficits

The forces behind internal monetary expansion in Pakistan can best be studied by analyzing the trend of credit creation separately for the private and the government sectors. The following table gives a breakdown of gross bank lending[1] by government and private sectors. The figures, which are on a July–June year basis, only broadly compare with the aggregate April–March year data in Tables I and II.

---

1. Gross bank lending (item 3 in Appendix Table I) is defined as the change in total internal assets of the banking system (including the Central Bank); the decrease in government cash balances is treated as indirect monetization and is shown on a net basis. Gross bank lending equals internal credit monetization plus the increase in banks' domestic liabilities other than money.

TABLE III

Gross Bank Lending

(Millions of rupees)

Years (July-June)	To Private Sector	To Government[1] Sector	Total
1951-52	231	462	693
1952-53	-122	616	494
1953-54	134	357	491
1954-55	-50	283	233
1955-56	100	161	261
1956-57	237	416	653
1957-58	42	679	721
1958-59	-4	190	186
Total	568	3164	3732

1. Net of aid counterpart funds.

Source: State Bank of Pakistan.

The above table shows that the bank lending to private and government sectors was marked by wide fluctuations from year to year. On the whole, however, net credit creation in the government sector, which we have defined as deficit financing in this study, has been the major factor behind monetary expansion. Lending to the government (including declines in government cash balances) accounted for 84 per cent of the total credit increase during the eight years, 1951-59. It is noteworthy that the level of deficit financing was at Rs. 1718 million during 1951-55, actually higher than the level (Rs. 1446 million) in the next four years.

Though our estimates of deficit financing have been arrived at indirectly[1] and refer to a slightly different period, they can be broadly related to the size of the over-all budget deficits and the level of development spending in the public sector.

1. The estimation of the deficit financing directly from the government budgets is not possible because the details of the government borrowing according to the source of borrowing are not being published at present.

	Millions of rupees	
	1951–55	1955–59
1. Public Development Spending	2281	3584
2. Over-all Budgetary Deficits[1]	1898	3698
3. Deficit Financing	1718	1446

It is clear that the size of the over-all budget deficit has been closely related to the rate of development spending in the public sector. The sharp increase in the budget deficits from Rs. 1898 million in 1951–55 to Rs. 3698 million in 1955–59 was attributable largely to the jump in the level of development spending in the public sector from Rs. 2281 million to Rs. 3584 million over the same period. It is also evident that ordinary revenue surpluses did not provide any more resources for the development sector in the later period than in the earlier years.

Roughly 79 per cent of the over-all budget deficit during 1951–55 was financed through the creation of money. That the monetary impact of the much larger budget deficit in the period 1955–59 was actually somewhat lower than in the 1951–55 period is explained by the fact that most of the increase in deficits between the two periods was financed by the larger inflow of foreign aid.[2] But it must be stressed that, even though foreign aid was financing an increasingly large portion of development spending in later years, deficit financing remained at a high level. Despite the decline in its relative importance, the creation of money financed over 40 per cent of the development spending in the public sector during 1955–59.

Private Sector Credit

In sharp contrast to the large expansionary influence of the government sector, private credit expansion was quite modest. In the four years 1951–55, bank lending to the private sector expanded on the whole by only Rs. 193 million, or barely 10 per cent of the total. The relative stability in the level of private bank credit was directly traceable to the large decline in

1. Covers both central and provincial governments' budgets. Deficits have been measured by total government borrowing (including borrowing from the central bank), drawal of cash balances, and foreign aid receipts.

2. Foreign aid helped finance the government deficit both directly and indirectly. Directly, a larger part of the budgetary expenditures was being financed through project aid and releases from the counterpart funds. Indirectly, the large net increase in the counterpart funds accumulated in the central bank against U.S. Commodity aid served as a powerful neutralizing influence on the central bank lending to the government. The net increase in counterpart aid funds was Rs. 669 million during 1955–59 as against nil in 1951–55.

the volume of foreign trade following the collapse of the Korean war boom and the tightening of import restrictions after 1952. In the period 1955–59, lending to the private sector expanded by Rs. 375 million, or little over 20 per cent of the total. A large part of this expansion occurred in credit to the manufacturing sector, which rose by Rs. 215 million; the share of the manufacturing sector in bank advances, which account for the bulk of commercial bank credit, increased from 24 per cent at the end of June 1955 to 35 per cent at the end of June 1959.

The relatively low bank lending to the private sector was against the background of an essentially passive monetary policy. There were no significant measures of quantitative credit control during the period under review. The legal reserve requirements for scheduled commercial banks, standing at 5 per cent of demand liabilities and 2 per cent of time liabilities, remained unchanged throughout. The Bank Rate, which had remained at 3 per cent ever since the establishment of the central bank in 1948, was raised only recently—to 4 per cent in January 1959. The only other monetary measures were in the field of selective credit controls; the most significant of these was the directive issued to commercial banks in June 1957 asking them to limit their advances against imported manufactured goods, bullion, foodgrains, and oil seeds to the extent of 60 per cent of the value of such goods, and also to limit unsecured advances.[1] The extremely limited monetary action to curb inflationary pressures may be explained largely by the fact that there are serious limits to the power of the central bank to control credit if the government operations are concurrently adding to the money supply.[2] If, as is the case in Pakistan, bank lending to the private sector accounts for a rather small portion of the total money supply, no amount of curtailment of credit to the private sector may be sufficient to offset the primary expansionary impact of a relatively large budget deficit. The task of controlling the secondary credit expansion was also not serious, because the fact that the ratio of currency to money is high in Pakistan (an average ratio of 2:3) itself ensures a low credit multiplier.[3] Furthermore, the decline in foreign exchange reserves in the earlier years also offset the expansion in primary reserves caused by government deficit finance operations.

---

1. These restrictions, which were withdrawin in late 1958, were mildly successful. Cf. State Bank of Pakistan, Department of Research, Report on Currency and Finance, 1958–59 (Karachi, 1959), p. 37.

2. See Joachim Ahrensdorf, "Central Bank Policies and Inflation" in International Monetary Fund Staff Papers, Vol. VII (October 1959), pp. 274–301.

3. The credit multiplier is defined as being equal to the reciprocal of $c + r(1 - c)$, where c stands for the ratio of currency in circulation to money and r for the legal or customary reserve ratio maintained by banks against deposits.

Accruals to Time Deposits

An important offset to the total credit creation in Pakistan has been the steady increase in non-monetary liabilities of the banking system, i.e., time deposits held with commercial banks.[1] The increase in these deposits absorbed 17 per cent of the gross credit creation during 1951–55 (Appendix Table I). During 1955–59, however, this percentage dropped to 11 per cent, indicating some shift in asset preference, probably not unrelated to the growing price pressures in the latter period. The ratio of time deposits with the commercial banks to their demand deposits, which had risen from 21.3 per cent in March 1951 to 40.0 per cent in March 1955, fell slightly to 39.9 per cent by March 1959. Even though the relative importance of time liabilities in money and quasi-money assets has not increased since 1955, their volume has been growing. Thus accruals to time deposits have remained an important leakage from the credit stream despite the drop in the value of money.

Liquidity Behavior and Price Developments

The wide swings in liquidity were a characteristic element in monetary developments in Pakistan during 1951–59. As Table IV shows, the monetary ratio (M/Y), which in 1951–52 stood at 0.192, rose steadily in the next few years, with the exception of 1952–53, to reach the peak of 0.255 in 1954–55. The net rise in the monetary ratio over the first period was of the order of 35 per cent. After 1954–55, the monetary ratio declined steadily every year and in 1958–59 was only marginally above the level in 1951–52. Admittedly, the data for money national income on the basis of which the monetary ratios have been computed is far from perfect. Still, there cannot be much doubt about the broad conclusion that income velocity 'fell' rather markedly in the years up to 1954–55 and 'rose' steadily thereafter, there being no significant change over the period as a whole.

The long-term change in the money-income ratio can significantly influence the monetary picture of a growing economy. Indeed, the rise in the demand for money at a rate faster than the rate of production growth is generally regarded as providing a powerful argument for the safe use of deficit financing in developing countries. The scope provided by the long-run increase in the monetary ratio for non-inflationary money creation can, however, be easily exaggerated. It is necessary to look at the whole question of the safe limits of deficit financing more closely. One of the primary reasons why the need for stocks of money may be expected to rise at a decidedly higher rate than money income in the early stages of development is that the importance of the monetary sector vis-à-vis the non-monetary sector, or subsistence sector, will tend to increase over time.[2] It is also, perhaps,

1. Includes saving deposits the withdrawal of which is subject to notice. The saving deposits not subject to notice are included under demand deposits.

2. If we assume that the share of the non-monetary sector stands at 1/3 of the total and that the entire future increases in output enter the monetary

TABLE IV

Money National Income, Money Supply and Monetary Ratios

Fiscal Year (April-March)	Money National Income (Y) (million Rs.)	Money Supply (M) (End of the Year, million Rs.)	Monetary Ratio[1] (M/Y)
1950-51	17,150	3290	0.192
1951-52	18,469	3549	0.192
1952-53	18,275	3315	0.181
1953-54	18,156	3693	0.203
1954-55	15,297	3915	0.225
1955-56	19,106	4533	0.237
1956-57	22,620	5043	0.222
1957-58	23,763	5303	0.223
1958-59	26,811	5451	0.203

Note: Money national income figures up to 1953-54 are Central Statistical Office estimates. Since then, national income figures have been published in constant prices. For the derivation of figures of money income from 1954-55 onwards, see Appendix. Money supply figures are taken from the State Bank of Pakistan Bulletin.

1. The monetary ratio as calculated here represents the ratio of money supply at the end of the year to money national income in that year. Strictly speaking, a monetary ratio based on annual averages of money supply should be used. This, however, can be done only at the cost of added complexities for the model employed.

(footnote continued from p. 488)

sector, it is evident that the marginal money-income ratio will be 50 per cent higher than the average money-income ratio, given a constant income velocity in the monetary sector.

inevitable that, in a country where capital and money markets are not well organized and well spread, a part of the savings should be held in the form of money hoards. Available studies of structural changes in income velocity further point out that the growing differentiation of production in the course of the development of the economy interrupts the synchronization of payments and calls for growing cash reserves.[1] Finally, hoarding of money balances may be regarded as a luxury and thus a function of the absolute level of income and wealth. The historical experience and the comparison between developed and less developed countries affirms a marked secular upward trend of the money-income ratio in a developing economy.[2] The effect of most of the above-mentioned factors tending to raise the money-income ratio—the Marshallian k— is, however, felt very slowly.[3] As we have seen, in the case of Pakistan the increase in the liquidity ratio over the eighty-year period was only 5 per cent. The net expansion in the money supply[4] absorbed by the rise in the liquidity ratio over this period was only Rs. 189 million. It also must not be forgotten that the "safe" limit of deficit financing in terms of the resource availabilities depends on the absolute level of k, which is often quite low in underdeveloped countries. Assume that the output in a less developed country is growing at an annual rate of 3 per cent and that the money supply can expand by 4.5 per cent annually without causing net inflationary pressures. If we further assume that the income velocity is 5—more or less the level in Pakistan—then the non-inflationary money expansion can at best equal 0.9 per cent of the national product. This, in turn, would, even with a relatively low credit[5] multiplier and the high ratio of currency to money, limit the safe annual deficit financing to only 0.4 per cent to 0.6 per cent of the national product.

---

1. H. S. Ellis, "Some Fundamentals in the Theory of Velocity," Quarterly Journal of Economics (May 1938), reprinted in American Economic Association, Readings in Monetary Theory (New York: The Blakiston Company, 1951).

2. E. M. Doblin, "The Ratio of Income to Money Supply: An International Survey," Review of Economics and Statistics, Vol. 33 (August 1951), p. 204.

3. For an interesting model of monetary analysis based on the assumption of a constant income velocity of money (unchanged money-income ratio), see J. J. Polak, "Monetary Analysis of Income Formation and Payments Problems," International Monetary Fund, Staff Papers, Vol. VI (November 1957), p. 1–50.

4. However, as we discussed, the time deposits expanded on the whole at a rate faster than the increase in the money supply and thus absorbed a part of the monetary expansion indirectly.

5. The higher the ratio of currency to money, given the deposit-reserve ratio, the greater would be the scope for credit creation by the government within the limits of total money expansion.

We shall now turn to a discussion of the significance of the rather sharp short-term changes in the liquidity ratio. It seems that the rise in the monetary ratio from 1952–53 to 1954–55 and its decline thereafter was in response to what may be termed autonomous changes in the price level. Between 1952–53 and 1954–55, the general price level, surprisingly enough, fell by 25 per cent, while it increased by 65 per cent in the next four years. Falling income velocity in periods of falling money incomes and price level and rising income velocity in periods of a rising price level and money income are generally regarded as a cyclical phenomena. However, it is possible to furnish a more concrete explanation for the inverse relationship of the short-term changes in the price level and the monetary ratio in an underdeveloped country. To the extent that the prices of agricultural crops produced primarily in the subsistence sector—foodgrains in Pakistan—fluctuate more widely than the general price level, the over-all movements in money national income do not represent adequately the changes in the value of output entering the monetary sector. If, as was the case in Pakistan during 1953–54 and 1954–55, the market prices of subsistence crops are falling sharply, the reduction in national income measured in money terms might be larger than the drop in the current value of the output entering the monetary sector. The decline in the monetary ratio, given the money supply, thus might appear to be greater than it actually is. The reverse can easily happen in a period of rising prices. To the extent that they reflect such changes, the movements in the monetary ratio are merely illusory.

The above factors, however, help to explain only part of the changes in income velocity in Pakistan. The increase in the monetary ratio after 1952–53 also probably reflected the building up of relatively large idle balances in the wake of tight import restrictions. The reduction in the availability of imported goods presumably led to a postponement of consumption and investment expenditures. The imports of consumer goods in 1953 and 1954 had been cut to an extremely low level. The cotton textiles and yarn imports, which had accounted for nearly 1/3 of private imports during 1951–52, were reduced to a negligible level in 1953. Though domestic production of cotton textiles was rising, it had not, as yet, compensated for the loss of imports. Moreover, the finer varieties of cloth were in scarce supply. The restricted imports of consumer and investment goods served as an important restraint on spending. The building up of cash balances and the forced saving implied by it, however, proved short-lived. By 1955, domestic industrial production had assumed significant proportions. The country had become self-sufficient in cotton textiles. The imports of consumer goods under the United States commodity aid program were further relieving shortages. The critical thing, however, was not the size of the consumer goods imports, which remained small, but the increased possibility of substitution between home and foreign goods due to the mushroom growth of import-competing production and the generally increased elasticity of supply of the new industrial sector. The postponed consumption expenditure therefore tended to be expanded at the cost of money holdings.

Another factor in the decline in the income velocity of money in the earlier period was probably the lags in spending out of the inflated profits. The profit incomes from trade and nascent manufacturing, which were in all likelihood limited up to 1952 due to the pressure of liberal imports, received a boost in the period from 1952 to 1955. The profits from the growing textile industry were particularly large. The almost sudden transformation of the profit picture tended to support the tendency toward hoarding in the short run. In contrast, with the growth of inflationary pressures after 1955–56, there was dishoarding by wide sections of the community. This dishoarding was not necessarily connected with dissaving from the point of view of individuals.[1] In the agricultural sector the farmers probably reacted by hoarding grain,[2] particularly in view of the pressing food shortage.

The general price level in Pakistan has been strongly influenced by the behavior of export prices and incomes. The drop of 44 per cent in export prices between 1951–52 and 1954–55 (see Table V), a result of the progress of recessionary influences in the world commodity markets and an equally large drop in export incomes, was the main factor behind the reduction in the general price level, despite net domestic monetary expansion. The crucial position of export prices and incomes in determining the general price level in a foreign trade-oriented economy like Pakistan was again demonstrated after 1955 when, following devaluation of the Pakistan rupee, higher export incomes (in local currency) helped reverse the downward trend in the price level. These autonomous changes somewhat limit the significance of the movements in the general price level as a guide to the inflationary impact of the internal monetary expansion. However, from 1956–57 onwards, internal monetary expansion seemed to have exerted a much stronger pull on the general price level. Export incomes and prices dropped once more between 1956–57 and 1958–59, but the internal price level rose by 18 per cent.

The over-all increase of 24 per cent in the domestic product price index between 1952–53 and 1958–59 may not appear to be excessive. However, the index as an indicator of the general price trends in the economy needs some qualification. The index is heavily weighted by prices of foodgrains, the production of which accounts for about 30 per cent of the total output. The prices of foodgrains were, to some extent, kept in check by measures of government control, in particular by compulsory procurement and restrictions on movement. The actual increase in domestic product prices therefore somewhat understates

---

1. In so far as an increasingly large portion of savings was probably being directed to gold and jewelry purchases, real estate speculation, and transfer of capital abroad, there was no direct effect on the internal price level as measured by the index of domestic product.

2. The amounts may be very small for individual farmers but, due to the very large number of farm households, total hoarding may be significant.

TABLE V

Price Indices
(1952-53 = 100)

Year	Domestically Produced	Export Goods	Import Goods	Terms of Trade
1951-52	105	155	119	129
1952-53	100	100	100	100
1953-54	93	82	104	79
1954-55	75	87	97	90
1955-56	94	99	135	73
1956-57	105	118	171	69
1957-58	110	121	184	66
1958-59	124	109	201	54

Source: Central Statistical Office for export prices, import prices, and terms of trade. See Appendix for Domestic Product Price Index.

the extent of the price pressures.[1]

Production Trends

The growth of production in the face of pressures of excess demand was small and absorbed, on the whole, only a very small portion of the total monetary financing (see line 5 of Tables I and II). The production increase has been particularly slow since 1954-55. The national income at factor cost, which rose by 15 per cent in the four years preceding 1954-55, rose by only 6 per cent over the following four-year period. Allowing for the effect of the sharp deterioration in the terms of trade, the increase in real output between 1954-55 and 1958-59 was of the order of only 4 per cent, which falls short of the estimated population increase of 5.4 per cent over the same period.

1. This implies that the net increase in the monetary ratio (decline in income velocity) is somewhat overstated. This strengthens our conclusion above. Also, to the extent that the real income increase is not adjusted for the terms of trade effect, the production growth is exaggerated.

Real National Income at Factor Cost[1]
(Prices: Average of 1949-50 to 1952-53)

(Million Rs.)

1949-50	17,236	1954-55	20,913
1950-51	18,140	1955-56	20,840
1951-52	18,066	1956-57	22,186
1952-53	18,796	1957-58	22,200
1953-54	20,024	1958-59	22,277

Source: Planning Commission.

The modest over-all increase in production, however, conceals a significant change in the structure of output. The period under consideration was characterized by a marked expansion in large-scale manufacturing, the income from which rose nearly fivefold—from Rs. 313 million in 1950-51 to Rs. 1,508 million in 1958-59; the share of large-scale industry in domestic production had risen to 7 per cent by 1958-59.

Index of Manufacturing Production[2]
(1954 = 100)

1951	48.6	1955	126.9
1952	58.6	1956	143.9
1953	76.3	1957	152.4
1954	100.0	1958	162.4

Source: Central Statistical Office.

The key development in the industrial field was the establishment of the textile industry, which in 1954 accounted for nearly 45 per cent of the value added in large scale manufacturing. The rapid growth of cotton textiles, accounting for 80 per cent of all textile output, is indicated by the following figures:

---

1. These estimates differ from the national income accounts of the Central Statistical Office in the following respects:
   (1) Terms of trade effects are excluded.
   (2) Depreciation is not deducted from estimates of gross value added.
   (3) The figures for minor agriculture crops, livestocks and fisheries, which are shown as constant in C.S.O. national income accounts after 1953-54, have been adjusted on the basis of later data collected by the Central Statistical Office.

2. Large-scale manufactures only.

	Capacity (at the end of the year)		Production	
Year	Spindles (Millions)	Looms (Thousands)	Cloth (Million Yards)	Surplus Yarn (Million Pounds)
1950	0.29	5.33	106.3	13.39
1954	1.32	18.42	345.2	99.95
1958	1.88	28.17	576.2	188.70

Source: Central Statistical Office.

The peak growth of the cotton textile industry was during the years 1952, 1953, and 1954, largely under the impact of the almost complete stoppage of cloth imports. The development of the textile industry not only helped to diversify the production structure and reduce dependence on imports, but it also was an important influence in cushioning price increases during 1955–59. Cotton textile prices, which had attained a high level in 1953, dropped almost steadily in the next three years, due mainly to increased competition among domestic producers since imports remained negligible. The continued decline in cotton textile prices in 1956, following the devaluation, was particularly healthy because it materialized notwithstanding an increase in the domestic price of raw cotton. In 1957, the trend of textile prices was reversed, but even so the rise was modest.

In sharp contrast to the rapidly expanding industrial production, agricultural output has tended to stagnate in the recent period and was a major factor in accentuating the imbalance between the aggregate demand and supply position. The production of major agricultural crops showed little net increase between 1954–55 and 1958–59, indicating a very low supply response to a marked increase in agricultural prices over the period. The unsatisfactory trend in agricultural production was only partly due to natural calamities like floods and droughts and was a reflection of the basic fact that the returns in agriculture have not been commensurate with the increase in the development expenditure. On the one hand, the gestation period of projects was too long, and on the other, the increase in irrigation facilities had not resulted in a corresponding increase in the area under cultivation because of the disappointingly slow progress on colonization schemes. [1]

Index of Agricultural Production[2]
(1949–50 to 1952–53 = 100)

1951–52	97	1955–56	99
1952–53	100	1956–57	113
1953–54	109	1957–58	110
1954–55	105	1958–59	107

Source: Central Statistical Office.

1. State Bank of Pakistan, Governor's Speech to the Ninth Annual General Meeting of the Shareholders (1957), p. 60.
2. Major agricultural crops only.

TABLE VI

Over-all Balance of Payments
Fiscal Year: April-March

(Millions of rupees)

	1951-52	1952-53	1953-54	1954-55	1955-56	1956-57	1957-58	1958-59
I. Imports of goods and services	2753	2389	1556	1609	1820	2848	2738	2687
II. Exports of goods and services	2718	1597	1481	1295	2008	1924	1780	1728
III. Excess imports	35	792	75	315	-188	926	958	959
Financed by:								
A. Net unilateral transfers from abroad	3	25	135	56	265	712	720	834
B. Decrease in foreign assets of the banking system (- increase)	42	676	-80	67	-411[a]	115	285	65
C. Decrease in net external assets of the non-bank sector	-5	45	25	208	42	39	-6	57
D. Suspense item and errors and omissions	-6	46	-2	-16	-84[a]	58	-42	4

a. Adjusted for revaluation of assets.

Source: Official balance of payments statements. For certain adjustments see Appendix.

The failure of production to rise rapidly affected not only the absorption of monetary financing directly, but also precluded the possibility of desirable liquidity changes. The stagnant agricultural production led, on the one hand, to the emergence of a lasting food deficit from 1955 onwards, and on the other, limited the availability of export crops, notably cotton. The large foodgrain imports and the reduction in the volume of exports were the two important causes of the growing balance of payments disequilibrium.

Balance of Payments Developments

Balance of payments difficulties dominated the economic scene in Pakistan almost throughout the period with which we are concerned. Strong payments pressures were exerted, on the one hand, by the growing internal monetary disequilibrium and, on the other, by the almost persistent deterioration in the terms of trade. On both these grounds pressures were cumulative and were intensified in the later years—1955–59. However, foreign aid imports, which were rather small up to 1954–55, rose sharply thereafter. Total aid-financed imports amounted to Rs. 2,531 million in 1955–59, as against only Rs. 219 million in 1951–55. Large aid availabilities, which represented a net addition to the country's resources, made possible the running of sizable import surpluses and the sustaining of development efforts. Whether they materially relieved the pressure on the balance of payments is a more doubtful matter.[1]

As we noted, a sizable part of the domestic monetary expansion in the earlier years was absorbed by the drop in foreign exchange reserves, which cushioned the impact of domestic inflationary forces. However, the level of foreign exchange reserves which came to prevail towards the end of 1952 was considered too low to permit further drawing down, even for short periods. The balance of payments policy, as reflected in the framing of foreign exchange budgets and administration of import control, aimed at a balance between foreign exchange earnings and outlays. The foreign exchange reserves showed significant improvement in 1955–56, in the wake of devaluation, but were quickly utilized in the next few years. For the period 1955–59 as a whole, there was little net change in the external assets of the banking system. Thus, in the period in which inflationary pressures really gathered momentum, the major instrument of policy for maintaining equilibrium in the balance of payments, in the extremely narrow sense of avoiding any net change in external reserves, was the change in the level of import restrictions. Even though the rupee was devalued by 30 per cent in August 1955,[2]

1. For assumptions necessary for the favorable effect on the balance of payments of a foreign loan and/or aid transfer, see J. J. Polak, "Balance of Payments Problems of Countries Reconstructing with the Help of Foreign Loans," Quarterly Journal of Economics (February 1943), reprinted in the American Economic Association, Readings in Theory of International Trade (Philadelphia: The Blakiston Company, 1949).

2. In September 1949, Pakistan was the only Sterling Area country which did not devalue her currency. Thus, the 1955 devaluation merely restored the pre-1949 par value of the Pakistan rupee vis-à-vis Sterling and the Indian rupee.

the reliance on the instrument of quantitative restrictions on imports and payments did not diminish. The devaluation failed to have a lasting beneficial effect on the balance of payments[1] because of the continued monetary expansion, the inelastic agricultural output, and the autonomous deterioration in the terms of trade.

II. Policy Instruments and Targets: Some Strategic Aspects

This section will attempt to relate the monetary developments in Pakistan discussed in the previous chapter to the real changes in saving and investment. The leverage of the instrument of credit creation on the level of the real capital formation will be discussed. The critical role of foreign exchange availabilities for investment will be stressed, and the reasons for the failure of exchange control as a tool for allocating resources under conditions of excess demand will be examined.

1. Deficit Finance and Capital Formation

The major goal of development policy in Pakistan has been the stepping up of the rate of capital formation, and this can be assumed to have been the main force behind the policy of deficit financing. In a sense, therefore, the success of the program of deficit financed development, quite apart from the social and economic consequences of inflation, may be judged from the extent to which it enabled the diversion of larger total real resources towards investment. The saving and investment flows in the Pakistan economy during the period 1951–59 are analyzed in this section. We shall assume in the following discussion that the monetary effects of the large foreign aid flows were neutral. Foreign aid represents a government to government transfer, and the assumption of a marginal propensity to spend of unity for the government, given the pressing need for development resources, appears reasonable. It will also be assumed that the deterioration in the terms of trade, with the low elasticity of substitution between home and foreign goods, contributed to the balance of payments pressures but did not influence investment more than consumption. It is not necessary for our purposes that these assumptions hold in the extreme sense. It can be readily conceded that the effect of the large foreign aid flow on availabilities was, in some sense, larger than the net transfers because foreign exchange was a scarce resource. But, on the other hand, the deterioration in the balance of payments resulting from adverse terms of trade probably did influence investment more than consumption because of the greater dependence of investment on capital goods imports.

Table VII shows the sources and the uses of investment funds for the periods 1951–55 and 1955–59. On the uses side, the estimates of public development outlays are based on central and provincial governments' capital budgets, while the gross private investment figures are based on the Planning

1. We are speaking of the improvement in an ex-ante sense, i.e., a reduction in degree of restrictiveness of imports and/or the price disparity between home and foreign price levels.

TABLE VII

Sources and Uses of Investment Funds (in current prices)
(Million Rs.)

	Sources	1951-52 to 1954-55	1955-56 to 1958-59		Uses	1951-52 to 1954-55	1955-56 to 1958-59
Domestic:	1. Gross Domestic Savings (Public and Private)[1]	3,656	3,925	7.	Public Development Expenditure[3]	2,281	3,584
Foreign[2]	2. Decline in the External Assets of the Banking System	705	55	8.	Gross Private Fixed Investment[4]	2,372	2,660
	3. Foreign Aid (gross)	219	2,531	9.	Changes in Stocks[5]	200	500
	4. Foreign Investment	70	58				
	5. Net Foreign Loans (including medium terms credits)	282	120				
	6. Change in Other External Assets and Liabilities	- 79	- 45				
	Totals	4,853	6,644			4,853	6,644

1. Gross domestic savings are estimated as a residue item. They include both private and public savings.

(Footnotes continued on following page)

2. Items (2-6) correspond to the excess imports (ajdusted for errors and omissions and suspense item) in the balance of payments statements. See Table VI. The foreign aid figures are gross receipts and include all loans payable in local currency. Insofar as part of the counterpart funds under U. S. aid are set apart for non-development purposes, the gross aid figures exaggerate the increase in availability of development resources and correspondingly show somewhat smaller allocation of domestic resources for development. This factor, however, was operative only in 1955-59. The sum of items (4), (5), and (6) above corresponds to the 'Decrease in external assets of the non-bank sectors' in Table VI.

3. The figures cover only those development outlays which are included in the capital budgets of the central and provincial governments. Some (probably 15-20 per cent) of the public expenditure on development is covered by the revenue budgets. The coverage of our figures, though not quite complete, is consistent over time. The public sector investment, according to the Second Five-Year Plan definition of development outlay, amounted to Rs. 4,270 million during 1955-59.

4. Planning Commission estimates. The figures cover fixed investment in industry, construction, and transport. The private investment in agriculture, rural housing, and other sectors is not covered because estimates are not possible. Also excluded are development expenditures by local bodies. For 1951-55 estimate, see Government of Pakistan Development Commission, The First Five Year Plan (Karachi: Manager, Government of Pakistan Press, 1955) p. 133.

Yearly breakdown for 1955-59 is as follows:

1955-56	Rs. 650 m	1957-58	Rs. 670 m
1956-57	Rs. 660 m	1958-59	Rs. 680 m

5. Figures of changes in stocks are rough estimates based largely on the Census of Manufacturing and statistics on the stocks of excisable commodities.

Commission estimates of fixed investment in construction, transport, and industry. The coverage of both our public and private investment is not complete because estimates of development spending in the revenue budget and of private investment in rural housing and agriculture are not available for the entire period. The narrow definition adopted here, which, roughly speaking covers, about 75% of all investment, has the advantage of consistency over time. On the sources side, items (2–6) correspond to the excess imports[1] in the balance of payments. Gross domestic saving is estimated as the balancing residual item and thus includes both private and public saving.[2] Crude though it may seem, this method offers the only practical basis for estimating savings in the absence of detailed national income accounts. One important qualification may be noted. To the extent that foreign aid receipts are shown on a gross basis, gross domestic savings are underestimated for the 1955–59 period. In this period, a substantial part of the counterpart funds accumulated under aid programs and held by the United States Government was diverted to non-development uses outside the government budget, such as United States military aid,[3] the United States Government expenditure connected with the administration of aid, etc. No direct estimate of these expenditures is possible, but the utilization of aid in the budget and net addition to counterpart funds over the period suggests that the leakage on this account from gross aid flow was probably not less than Rs. 800–900 million.[4]

Despite shortcomings and incomplete coverage in some respects, the saving and investment accounts as presented here can be helpful in a broad comparison of the changes in the real capital formation between 1951–55 and 1955–59. However, we first have to solve the somewhat intricate problem of a suitable price deflator for gross investment in money terms. The general price index relating to domestic product rose from an average of 93.2 (1952–53 = 100) during 1951–55 to 108.2 during 1955–59—a rise of 16 per cent.[5] Over the same period, the import price index jumped up from 105.0 (1952–53 = 100) to 172.7, depicting a rise of about 64 per cent, mainly reflecting the effect of devaluation. In

---

1. Adjusted for the errors and omissions and suspense item. See Table VI.

2. Public savings amounted to Rs. 383 million during 1951–55 but were negative by Rs. 114 million during 1955–59. However, if it was possible to include the development expenditure in the revenue budgets under public investment, public savings would be correspondingly higher.

3. The expenditure on this account was estimated at Rs. 90 million and Rs. 123 million in 1957–58 and 1958–59, respectively, in the Finance Minister's Budget Speech, Government of Pakistan Budget, 1958–59, p. 12.

4. Taking account of this would raise the gross domestic saving figure for 1955–59 Rs. 4,725 million as against Rs. 3,656 million in 1951–55.

5. The price increases both in the general price index and import price index have been calculated on the basis of the arithmetic means of yearly indices. The price deflator based on these, therefore, is only a rough indicator. See Table V.

estimating the price increase relative to investment outlay, we have assumed an import coefficient[1] of 0.40. On this basis, the increase in the price of investment goods during 1955–59, as compared to 1951–55, works out at 35 per cent. Admittedly, this estimate is quite rough, but the conclusions that follow are sufficiently clear-cut to allow for a reasonable margin of error in either direction.

Adjusting the 1955–59 investment figures for price change, we find that the gross investment actually showed only a nominal increase of 1.5 per cent over 1951–55. Though the gross investment showed a small increase, net capital formation between the two periods can be taken to have declined substantially due to higher depreciation charges (on a rising capital stock) in the later period. The reduction in real net domestic investment during 1955–59 took place notwithstanding an increased utilization of net foreign resources. During 1951–55, a substantial part (nearly 1/4) of domestic investment was financed by drawing down foreign exchange reserves, foreign loans, and foreign aid. In the next four years, very little foreign exchange reserves were used, but foreign aid increased enormously with the result that, even after allowing for net internal transfers to non-development purposes from foreign aid and adjusting for the price increases, the real value of net foreign exchange utilizations exceeded the corresponding external financing in the previous period by 15–20 per cent. This means that the domestic resources allocated to development fell, between 1951–55 and 1955–59, even more sharply than the level of real net domestic capital formation.

The disappointing behavior of real capital formation in what were the first four years of Pakistan's First Five Year Plan—a serious attempt at comprehensive planning—at first seems odd; the rate of gross fixed investment was only 7 per cent of national income, though it must be emphasized that the coverage of our investment figures is not complete. The conclusion about the decline in investment in recent years appears incongruous with the increasing tempo of development programs in the public sector. But there is no real contradiction here. While the real investment in the public sector did rise (about 16 per cent in gross terms), this increase was more than offset by a sharp decline in private investment. Gross private (fixed) investment at Rs. 2660 million during 1955–59, though somewhat higher in money terms than the level in 1951–55, showed a decline of 17 per cent in real terms and, making allowance for increasing depreciation, at markedly higher replacement cost, was probably not much higher than 2/3 of the level in 1951–55; the share of private investment in the total fell from 50 per cent to 40 per cent over the period.

---

1. The First Five Year Plan estimated the import coefficient for the public sector program at 0.42 and 0.52 for private sector investment. We have chosen a somewhat lower coefficient partly because of the dominance of the public sector outlay in actual outlays and partly because of possible substitution in favor of home goods due to a rather marked divergence between prices of the latter and those of imported goods.

The reduction in private investment cannot be attributed to any lack of investment demand as such.[1] The available evidence suggests that profits were being maintained at a high level. According to the Census of Manufacturing, gross profits (value added minus employment costs) as a per cent of sales (value of products made) dropped only slightly, from 23 per cent in 1954 to 20 per cent in 1957. Gross profits in cotton textiles declined rather sharply from the abnormal level of 35 per cent to 24 per cent over the period, but in the rest of the manufacturing sector, which was expanding rapidly, profits remained unchanged. Even in the case of cotton textiles, manufacturers were exerting pressure on the government for liberal imports of machinery and spare parts, and the export bonus scheme,[2] first applied to the export of cotton manufactures in 1958, to some extent represented an accomodation of this pressure.

The most important single factor limiting the level of private investment appeared to be the scarcity of foreign exchange resources. The limitation of foreign exchange was critical because the import coefficient in the investment program is very high, and the substitutability between home and foreign goods in investment is also very limited. The frustrated investment demand implied that the savings (ex-ante) in the private sector could not be channeled into productive uses.[3]

An idea of the effect of import restrictions on the level of capital goods imports can be had from the following figures. Private capital goods imports were running at the annual rate of about Rs. 209 million during 1955–58 as against Rs. 177 million in the preceding four years. However, allowing for the effect of the devaluation in August 1955 and the rise in the import prices, the investment goods imports in real terms dropped by at least 25 per cent between the two periods. The actual decline in the availabilities of capital goods imports to the private sector was even larger because Pakistan Industrial Development Corporation imports (on account of the public sector industrial development program), included under private imports, in fact, expanded.

The reduction in foreign exchange allocation for private sector investment took place in spite of the high priority on capital goods imports and the existence of strict exchange and trade controls. Apparently, given the total foreign

1. The investment demand in an inflationary situation may actually be unduly inflated due to high profit margins and expectations of rising prices. Even if we abstract from the effect of price expectations, the normal investment demand can in such a case be determined only by using a shadow price for the exchange rate. We are assuming that it was this normal investment demand in Pakistan which was frustrated due to foreign exchange difficulties.

2. Under this scheme, a part of the foreign exchange earned could be utilized directly by the exporter for imports of specified items.

3. The unsatisfied investment demand was probably directed to real estate and stock market speculation as well as to capital transfers abroad. The real balance effect operating through the rise in price also absorbed a large part of ex-ante savings by increasing the need for cash balances.

exchange availabilities, the substitution in favor of capital goods as against consumer goods and raw materials just could not be pushed further. In fact, the inadequate imports of raw materials and spare parts had already resulted in substantial idle capacity.

The fact that the limit on private investment, and thus on total investment, seemed to operate from the balance of payments side should not mislead us. On the assumption that the terms of trade deterioration and large aid inflows together had only neutral effects on the payments position, the balance of payments weakness can be largely attributed to internal monetary pressures. The real phenomenon was that, after 1954–55, consumption was not being reduced to release resources in response to the deficit financed development spending by the Government. The real consumption demand appeared to be relatively inflexible, so that the attempt by the Government to raise the level of capital formation in the public sector succeeded, after a point, only at the cost of a curtailment in private investment activity.

2. Exchange Control and Resource Allocation

We argued in the preceding section that the serious foreign exchange shortages, which appeared to be a limiting factor in private investment, were largely a reflection of the growing internal monetary disequilibrium and therefore the balance of payments problem, though serious, was secondary. However, the mechanism through which excess demand forces operated on the balance of payments and made a larger leakage to consumption possible, despite stringent exchange control, is not without interest. The study of the forces limiting the effectiveness of exchange control as a tool for allocating scarce foreign exchange resources for investment also helps to bring out some of the factors sustaining real consumption. The behavior of real consumption in Pakistan under conditions of excess demand—given exchange control—will be analyzed in this section with special reference to the emergence of food shortages and the decline in foreign exchange earnings, which were the two main elements in the balance of payments deterioration after 1955. While the drop in earnings served to limit the import capacity, the overriding need for foodgrains imports constituted a direct leakage from the resources subject to the administration of foreign exchange control.

However, before we turn to these specific factors affecting the operation of foreign exchange control in Pakistan, it would be helpful to outline the basic relationships between the twin policy instruments of deficit finance and exchange control, on the one hand, and the rate of spending on the other.

Credit Creation and Exchange Control

The use of the instrument of deficit finance and the resulting inflationary pressures almost inevitably give rise to a balance of payments problem for a country.[1] The payments problem will emerge even if the internal rise in

1. "In a real sense the credit expansion is the cause of payment problems." J. J. Polak, "Monetary Analysis of Income Formation and Payments

prices is relatively successful in inducing forced savings. In that case, price effects operating through the price elasticities may dominate over the income effects and the marginal propensity to import. The extent to which balance of payments pressures can be allowed to result in actual deficits is often seriously limited by the availability of foreign exchange reserves and/or foreign borrowing. The use of the instrument of monetary expansion for raising the level of capital formation therefore necessitates additional instruments, the most frequent being the use of quantitative restrictions and exchange control for achieving the target of 'equilibrium' in the balance of payments. The corrective balance of payments policies, however, often conflict with the attempts to raise real investment expenditures, which may be at the back of the forces inducing monetary expansion. This point needs some elucidation.

The modern theory of balance of payments[1]—the so-called absorption approach—emphasizes that, excluding the transactions connected with external credit monetization, the balance of payments can be regarded as a difference between the value of a country's output and its expenditure. If the balance of payments equilibrium, in the narrow sense of no net change in international reserves, is to be maintained, additional net real expenditure desired by the government must be provided for either by 1) a reduction in expenditure in other sectors, or by 2) an increase in output in excess of the induced increase in expenditure, or 3) both. The possibilities of a marked increase in production in the short run are generally limited. Thus, imposition of import restrictions and administration of exchange controls can, given the relative inelasticity of output, help to maintain foreign balance only by reducing real expenditure indirectly through an increase in the price level and/or by the forced postponement of outlays due to the non-availability of scarce foreign exchange resources. In so far as the primary objective of the autonomous increase in the government outlay, which generates the excess demand, may be the increase in total investment, consistency further demands that the balance of payments policy, designed to keep receipts and payments in balance, should operate mainly through a reduction in consumption expenditures. That the squeezing of real consumption may not always be easy in view of the structural characteristics of an underdeveloped economy will become clear in the following discussion.

Excess Demand: Marketable Surplus and Foodgrains Imports

Pakistan, which up to 1954 was self-sufficient in foodgrains, at least in normal crop years, has, for the last four or five years, been faced with a

---

Problems," I. M. F. Staff Papers (November 1957). "A continuing balance of payments deficit ultimately requires credit creation to keep it going." H. G. Johnson, "Towards a General Theory of the Balance of Payments," in his International Trade and Economic Growth (London: Allen and Unwin, 1958), p. 157.

1. For an excellent treatment, see Johnson, ibid., pp. 153-168.

serious food shortage which has entailed heavy food imports. Foreign exchange expenditure on foodgrains, which was nil in 1955, has since 1956 become the single most important drain on the balance of payments. The value of foodgrain imports in the three years 1956 to 1958 exceeded the value of total capital goods imports. Nowhere is the claim of consumption on foreign exchange resources so clearly in evidence as in the case of foodgrains.

	Foodgrains Imports[1] (1)	Total Imports[2] (2)	(1) as per cent of (2)
	(Millions of rupees)		
1956	390	1986	19.7
1957	546	2096	26.0
1958	437	1888	23.1

1. Source: Ministry of Food and Agriculture.
2. Source: Central Statistical Office

The trend in foodgrain consumption, which accounts for perhaps 25 per cent of the total national expenditure, may be studied with a view to throwing light on the nature of the forces sustaining real consumption. A look at the figures of food crop production[1] makes it clear that the long-run factor of population growth occurring at a rate faster than the increase in foodgrains output has undoubtedly been a major influence in the foodgrains deficit. As against the steady rise in population, foodgrains production had risen by little more than 1 per cent between 1951-55 and 1955-59. However, in the last few years per capita foodgrains consumption has risen, and apparently at a rate faster than income per capita. Per capita consumption of foodgrains at 17.4 ounces per day during 1957-58 was 8 per cent higher than in 1954-55;[2] the income per capita was only 2-3 per cent higher over the period. Separate analysis of the food position of West Pakistan further highlights the expansion in foodgrains consumption and helps to pinpoint the role of excess demand in this regard. There was a marked increase in the consumption of wheat and rice in West Pakistan between 1955 and 1959, while the consumption (and production) of coarser foodgrains remained almost unchanged. Taking all foodgrains together, consumption in West Pakistan during 1959 was 20-25 per cent higher than in 1955, and was far in excess of either the population increase or the rise in per capita income.

---

1. Index of Food Crop Production (1949-50 to 1952-53 = 100)

1953-54	113	1956-57	112
1954-55	103	1957-58	106
1955-56	93	1958-59	101

2. 1954-55 (July-June) was the last year in which no foodgrains were imported. 1954-55 was also the year in which internal rice and wheat prices were at the lowest level since 1950-51.

The increased consumption and net imports of wheat in West Pakistan were in part a reflection of the increased quantity required by the government for controlled distribution—mostly in the urban sector. The wheat distributed through the government (procurement + imports – change in stocks) has averaged over one million tons annually in the three years following 1957, as against about 700,000 tons in the preceding three years, notwithstanding the fact that prices have been gradually raised in the controlled distribution areas since 1955–56.

In contrast with the growth in the urban sector requirements, and despite the steady increase in production as well as successive raising of the procurement price, the amounts of wheat procured by the government, which may be regarded as an indicator of the marketable surplus, have not risen. The total quantity of wheat procured by the government during 1957–59 was only 3 per cent higher than for 1954–56, while production over the period was 10 per cent higher. The market price of wheat rose rather sharply after 1955. But insofar as the procurement price was raised as much or more, the incentive to sell to the government was not affected.

Forced Saving and Structural Factors

The evidence is therefore strong that the food shortage, which was causing serious balance of payment difficulties after 1955, was traceable at least in part to the failure of the marketable surplus of foodgrains to increase consistent with the growing requirements of the non-agricultural sector. It is, however, clear that the foodgrain requirements of the newly employed workers need not, by itself, give rise to a net consumption pressure in the economy. At the root of the problem was the fact that resort was had to created money to finance new development projects which were not as yet adding significantly to output flow. The failure of the marketable surplus to rise merely signified that the foodgrain consumption in the subsistence agricultural sector was not being squeezed sufficiently to accommodate excess demand pressure arising elsewhere. Looked at in another way, the consumer spending generated in the monetary sector seemed to call for a direct addition to the wage goods supplies. The workers who were transferred to the development projects had to be fed and clothed, and their requirements represented, by and large, a net addition to the total 'real' consumption demand. It was probably not possible to squeeze real wages to any significant extent because it is perhaps the real and not the money wage which tends to regulate the flow of surplus workers from the agricultural sector to the non-agricultural sector. The concern with the real wage rate may have been at the back of the government effort to maintain reasonable prices of foodgrains through supplementing imports with internal controls and food subsidies.[1]

1. The sale of foodstuffs involved a subsidy of Rs. 152.4 million during 1956–57, Rs. 128.3 million in 1957–58, and Rs. 67.1 million during the 15-month period ending June 1959. The subsidy, which served to cushion the impact of growing food shortages and rising foodgrain prices on the urban population, involved a net drain on government revenue and thus further aggravated the monetary situation.

The above discussion, though confined to foodgrains, seems to suggest general difficulties facing the attempts to generate real savings through inflationary financing in less developed economies which have a large subsistence sector. The successful use of the instrument of deficit financing for raising the level of investment ultimately requires the curtailment of real consumption at a given income level. This may not be easy in economies where a substantial part of the transactions lies outside the range of monetary influences. As we pointed out earlier, nearly 30 per cent of the national income in Pakistan originates in the non-monetary sector. The consumption in this sector cannot be reduced by the simple expedient of bidding for real resources by the creation of money in the government sector.

Import Substitution: Internal Inflation and Exports

We will now discuss the difficulties of restraining consumption and making exchange control effective, with reference to the other important elements in the balance of payment situation in Pakistan, viz., the reduction in foreign exchange earnings. As the following table shows, the almost steady drop in exports after 1951-52 brought the receipts down to the record low level of $321.3 million in 1958-59. As a percentage of money national income, export earnings dropped from 7.3 per cent in 1952-53 to 5.7 per cent in 1958-59.

Export Receipts
(In Current Prices)

Years (July-June)	Million $	Years (July-June)	Million $
1951-52	662.7	1955-56	406.7
1952-53	405.3	1956-57	357.2
1953-54	393.5	1957-58	321.0
1954-55	367.3	1958-59	321.3

Source: Balance of Payments Records

The decline in exports has been due, apart from the reduction in export prices, to two main factors: a) large-scale import substitution as a direct consequence of the growth of industrial capacity, as typified by the newly established cotton textile industry, and b) the presence of inflationary pressures in the economy. The import substitution was a deliberate policy objective and was a corollary of the efforts to diversify production structure and to reduce dependence on the extremely vulnerable foreign trade sector. The success of the policy of import substitution was indicated by the fact that, starting from a position of almost complete dependence on foreign textiles, the country had already become self-sufficient in most varieties of cotton cloth and yarn by 1955.

The growth of industrial capacity, which has made import substitution possible, has adversely influenced the potency of import restrictions as a policy instrument in two ways. First, it has tended to lower the long-run equilibrium level of imports and exports, and has thus reduced the range of resources subject to the administration of the exchange control authority. Second, it has made exports susceptible to the domestic monetary situation. The setting up

of industry in Pakistan has meant a greater degree of substitution both between home products and foreign trade goods and between imported goods and export products. A necessary assumption for the effectiveness of the import restrictions under domestic inflationary conditions is that the rise in demand (and/or prices) of both home goods and imported products does not result either in a marked reduction of export supply (as distinct from the supply of export products) or in a lowering of the competitiveness of exports. In Pakistan, however, with the gradual change in the structure of the economy, exports were becoming increasingly sensitive to domestic inflationary pressures.

The internal inflation tended to affect the export earnings, not so much through a loss in the competitiveness of exports as through a reduction in the export supply. In a country like Pakistan, whose main exports consist of agricultural products, cost competitiveness in the normal sense has limited significance. The prices of Pakistan's major exports are internationally determined, while the supply of export products is not greatly sensitive to changes in cost conditions.[1]

The major instance of the reduction in export supplies due to internal demand pressures was the drop in raw cotton exports after 1954-55; the fall in the previous years represented largely a replacement of imports. The increase in the domestic absorption of raw cotton was the major factor in the marked drop in receipts from raw cotton from $104.6 million in 1954-55 to $46.7 million in 1958-59. The appearance of cotton manufactures exports, which stood at $14.3 million in 1958-59, only partly offset this decline.

Position of Raw Cotton
(Thousands of Bales)

Years	Production (Sept.-August)	Domestic Consumption (July-June)	Exports (July-June)
1954-55	1602	800	723
1955-56	1740	810	948
1956-57	1600	1000	664
1957-58	1700	1120	447
1958-59	1520	1100	460

1 bale = 392 lbs.
Source: Reports on Currency and Finance

The domestic cloth consumption,[2] which since 1954 has been met almost

1. With the small peasant farm as the unit, the costs of production consist essentially of the living expenses of the farmer and his family and of certain fixed dues to the landlord or the state. The high ratio of these supplementary costs to the fixed costs tends to keep production at all times more or less at the maximum level possible under the given conditions. K. N. Raj, The Monetary Policy of the Reserve Bank of India (Bombay, 1948), p. 43.

2. Estimated as follows: Domestic production of cloth and yarn minus net exports of cloth and yarn. 1 lb. yarn = 4 yards of cloth.

wholly from domestic sources, has risen rapidly. The estimated consumption of cotton cloth of 937 million yards during 1954 already compared favorably with the average of 868 yards in the three years from 1950 to 1952, the period of liberal imports. By 1958, cloth consumption was at a record level of 1306 million yards, and the per capita consumption at 15.3 yards already exceeded the First Plan's target of 14 yards for 1960 and compared very favorably with the average of 11.3 for the years 1950–52. The per capita real income during 1958–59 was barely 4 per cent higher than the average in the years 1950–53. The phenomenal growth in cotton cloth consumption, which seriously affected export earnings and thus import capacity, was fed by the presence of excess demand in the economy and was supported by the improvement in the relative prices of cloth vis-à-vis other commodities. The case of availability may have further shifted consumer expenditure towards textiles. That the steady increase in consumption of cloth was accompanied, at least until 1956, by a decline in prices of cotton textiles suggests that the income mechanism emphasized by the modern theory, rather than the price mechanism of the classical theory, was the dominant factor in the balance of payments adjustment.

It may be mentioned that the leakage from the export stream under inflationary pressures was not attributable to the growth of industrial capacity alone. The tea exports from Pakistan furnish a good example of the direct substitution in consumption between exports and non-exports. While annual tea production dropped slightly from the average of 53 million lbs. in 1951–55 to 51.5 million lbs. in 1955–59, the exports of tea dropped from an annual average of 27.0 million lbs. to 14.25 million lbs. between the two periods.

The forces operating on tea and cloth consumption were, to some extent, the same as were influencing foodgrains consumption. In one case, the response to the emergence of excess demand was the increase in imports, while in the other, it took the form of a reduction in exports. The process in the latter case was conditioned by the diversification of the economy and the growth of an industrial base. The growth was, however, largely in the consumer goods industries, so that the dependence of the investment program on capital goods imports remained heavy. This, in turn, operated against the attempts of the private investors to obtain real investment goods notwithstanding exchange control.

III. Conclusion

From the point of view of their impact, the deficit financing operations in Pakistan can be seen as falling into two distinct periods, 1951–55 and 1955–59. The level of money creation in the government sector did not show much change between these two periods. It was accompanied, however, by vastly different results. Stated somewhat crudely, deficit financing operations in the earlier four-year period can be judged to have been successful, at least partly, while their continuance thereafter was definitely harmful for the economy. The widely divergent experiences with a more or less given amount of money creation appears at first sight to be the most interesting feature of Pakistan's experiment in the use of monetary instruments for financing development.

The analysis shows that the conditions for deficit financing were not unfavorable in the earlier years of Pakistan's development effort, though this coincidence was probably a matter of accident rather than design.[1] The foreign exchange reserves had reached a record high level by mid-1951 and were to make possible a very sizable net withdrawal in the subsequent period. The absorption of a significant portion of the internal monetary expansion by large balance of payments deficits limited the average annual increase in the money supply to less than five per cent during 1951–55. As against this, the food situation during this period was, on the whole, quite satisfactory. The recovery after the wheat harvest failures of 1952–53 was rapid, and total foodgrain production in 1953–54 set a record which was not surpassed until 1959–60. At the same time, capacity in the consumer goods industries, particularly cotton textiles, was being rapidly built up and was providing the basis for a marked increase in industrial output.

That there was no serious imbalance between money demand and goods availabilities in the first period is evident from the fact that the general price level declined rather sharply. The significance of this decline should not, however, be exaggerated. The general price level in a foreign trade–oriented economy like Pakistan is greatly influenced by the behavior of export prices and incomes. The drop of 44 per cent in export prices between 1951–52 and 1954–55—a result of the progress of recessionary influences in the world commodity markets and an equally sharp decline in export incomes—was the main factor behind the reduction in the general price level during this period. The impact of these autonomous changes somewhat limits the significance of the movements in the general price level as a guide to the inflationary impact of the internal monetary expansion.

Another important element which appeared to cushion the expansionary impact of credit creation during 1951–55 was the sharp rise in the monetary ratio (the inverse of the income velocity of money). A large part of this increase did not, however, prove to be lasting. To the extent that liquidity increased due to a temporary postponement of consumption and investment expenditures, as a result of a sudden and drastic cut in imports and possible lags in spending out of inflated profits, the impact of money creation was merely shifted to the next period.

It is clear, however, that, on the whole, deficit financing operations posed no real threat to financial stability up to 1954–55. The balance of payments problem was emerging mainly due to the sharply declining terms of trade,

---

1. If more attention had been given to circumstances governing the success of the instrument of deficit financing, greater restraint and more flexibility in government tax and expenditure policies would have been shown after 1955, when strong price pressures had already emerged and a serious food crisis was threatening the economy. In fact, the peak levels of deficit spending were reached in 1956–57 and 1957–58, when the situation was already threatening to get out of hand.

but had not yet assumed serious proportions. Under the circumstances, the financing of public sector development outlays with created money definitely contributed to the mobilization of resources for capital formation; although the figures for private investment before 1951-52 are not available, it seems fairly clear that the stepping up of development spending in the government sector was not hampering private investment to any serious extent prior to 1954-55. But we must hasten to add that the magnitude of the resources thus mobilized was not large. A substantial part of the higher level of domestic investment during 1951-55 was accompanied by disinvestment in the foreign sector. Had large foreign exchange reserves not been available, the balance of payments pressures, which were to seriously impede the success of the deficit financed development program during 1955-59, could not have been accommodated and development imports would not have been available.

The deficit financing operations after 1955-56 were definitely inflationary. The net result of the growing inflationary pressures and accompanying foreign exchange difficulties was to frustrate the efforts of the government to obtain additional real resources for capital formation through created money. The maintenance of development outlays, financed through borrowing from the banking system during 1955-59, was largely at the cost of a sharp curtailment of private investment activity, resulting largely from the scarcity of foreign exchange resources. The limitation of foreign exchange was critical because the import coefficient for the investment program was very high and the substitutability between home and foreign goods in investment was also very limited.

It is not possible to establish a precise quantitative relationship between the mobilization of resources for the public sector and the frustration of investment demand in the private sector. It is significant, however, that the statistical evidence presented in this study strongly suggests that gross domestic savings, in real terms, fell between 1951-55 and 1955-59. The ratio of gross domestic savings to money national income, at 5.1 per cent during 1955-59,[1] also stood fractionally lower than the 1951-55 level (5.2 per cent). The rise in the share of consumption in the national product may be explained in part by the autonomous factor of a rate of population growth in excess of production increase. Nevertheless, the conclusion appears to be valid that the deficit-financed development program during 1955-59 was able to bring about little net diversion of real resources towards investment.

The Pakistan experience in the years after 1955-56 throws some useful

---

1. According to the Second Five Year Plan, the ratio of gross domestic savings to money national income during this period averaged 5.9 per cent; the figures for the earlier period are not given. It may be noted that the Second Plan employs a somewhat wider definition of saving and investment than is used here. The Plan estimates of money national income also differ in some respects from those worked out in this study. Cf. Government of Pakistan Planning Commission, The Second Five Year Plan (Karachi: Manager, Government of Pakistan

light on the broad issue as to whether inflationary financing, distinct from creation of money as such, can be successful in mobilizing real resources for development in underdeveloped economies which are almost entirely dependent on capital goods imports for their investment program. Our analysis indicates that much hinges on the strength of the balance of payments. If internal monetary imbalance impinges on an inherently weak payments situation, resulting, for instance, from secularly declining terms of trade, foreign exchange would become a major bottle-neck. The scarcity of foreign exchange resources would, in turn, tend to put an indirect brake on investment. The crucial point is that in countries where the import coefficient of the investment program is high, balance of payments difficulties tend, ceteris paribus, to frustrate investment decisions more than consumption demand. No doubt the authorities can, and do, supplement the use of the instrument of monetary expansion with the use of quantitative restrictions and exchange control for controlling the balance of payments and for bringing about the desired allocation of foreign exchange resources for investment. But the relevant question is: can the effectiveness of the exchange control tool under excess demand conditions be taken for granted?

In Pakistan's case, at least, excess demand forces operating on the balance of payments made possible a larger leakage to consumption despite stringent exchange control. Several factors combined to limit the efficacy of import restrictions as a policy instrument. The rapid growth of industrial capacity and resulting large-scale import substitution, which had been achieved by 1955, tended to lower the long-run equilibrium level of imports and exports and thus reduced the amount of resources subject to the administration of the exchange control authority. The diversified production structure also made exports more sensitive to domestic inflationary pressures. Furthermore, the existence of a large subsistence sector which was, so to speak, outside the range of monetary forces acted against an increase in the marketable surplus of foodgrains, which was needed to feed the newly employed workers. This aggravated the food shortage and resulted in additional claims on foreign exchange resources.[1]

Press, 1960), Table II, p. 28.

1. It seems likely that, given the dependence of the country on foodgrain imports, the total balance of payments effect of the public sector development program was not lower than the private investment which was foregone. The import coefficient of the public sector investment is supposedly lower, but the employment effect of government spending is greater. A part of the demand for wage-goods, which directly influenced the payments position in Pakistan, thus probably stemmed from increased development outlays in the public sector. This may also help to explain why the real consumption of foodgrains and cotton cloth appeared to be exceptionally buoyant during 1955–59.

In the final analysis, the failure of deficit financing policies in Pakistan to raise the level of real capital formation appreciably, if at all, must be attributed to structural factors. It may be argued that the raising of investment in the public sector was itself a policy goal. It is, however, difficult to accept the notion that the change in the structure of investment resulting from deficit financing operations was desirable. The preponderance of the public sector outlays led to an undue emphasis on investment in public utilities, which have particularly long gestation periods. Within the private sector, residential construction, particularly luxury housing, was probably unduly stimulated due to the general shortage of capital goods imports and speculative building. This distortion of the investment structure was already hampering the growth of production.

Appendix

Notes on Monetary and Balance of Payments Data

Monetary Data (Appendix Table I

The annual figures for internal credit monetization (line 2) have been arrived at by adding the change in the money supply to the net decline in the external assets of the banking system. The changes in foreign assets of the banking system are derived from the official balance of payments statements (text Table VI) rather than from the balance sheets of the central and commercial banks, which do not, in some cases, report foreign assets and liabilities separately. The source of the money supply data is the State Bank of Pakistan.

Changes in domestic liabilities other than money (line 4) represent only changes in time liabilities of the commercial banks. Gross Bank lending (line 3) is thus obtained directly by substracting line 4 from line 2.

External credit monetization is derived from the balance of payments statements (Table VI). It represents the sum of net unilateral transfers and the decrease in net external assets of the non-bank sector.

The calculation of the monetary ratio, on the basis of which liquidity changes are determined, posed certain problems. No continuous series of money national income for Pakistan was available and such a series had to be constructed. The Central Statistical Office, which has the prime responsibility for compiling the national income data, has, since 1954–55, published national income statistics in real terms only—in constant prices, average of 1949–50 to 1952–53. The task was, therefore, to convert the real national income figures from 1954–55 onwards to a series in current prices. This was made possible largely by the availability to the author of a Domestic Product Price Index prepared at the Institute of Development Economics, Karachi. The index given below furnished an estimate of year to year changes in general prices from 1953–54 onwards, which were combined with the real national income data and the basic money national income figures for 1953–54 to obtain a continuous current price series.

(1951–52 = 100)

1953–54	88	1956–57	99
1954–55	71	1957–58	114
1955–56	89	1958–59	117

The money national income series and Domestic Product Index figures thus obtained are given in text Tables IV and V and have been used for calculating line 7 in Appendix Table I. Our composite series are necessarily rough and should be so treated.

Balance of Payments Estimates (Text Table VI)

The balance of payments estimates contained in Table VI have been derived largely from the official balance of payments statements published by the State Bank of Pakistan up to the middle of 1960. However, to the extent possible, adjustments have been made for incomplete coverage.

In the official balance of payments statements, the coverage of imports under foreign aid and loans is not complete for the period up to mid-1955. An attempt has been made to meet the gap by independent estimates. The figures of foreign aid for the four years 1951–52 to 1954–55 have been arrived at by combining foodgrains imports under foreign aid[1] with the foreign aid figures in the central government budget.

The foreign loans (World Bank, loans from the U. S. Export-Import Bank, U. K. credit and medium term credits for capital goods) are not reflected in the official statistics up to 1954. The details of adjustments made under non-bank assets are given below:

1952–53	IBRD (World Bank) loans	Rs. 21.7 m
	US Ex–Im Bank loan	Rs. 49.6 m
1953–54	IBRD loans	Rs. 26.0 m

1954–55 Rs. 120 million — represents an estimate of net utilization of deferred payments arrangements. The imports against medium term credits contracted in that year are reflected in the official statistics at a time when installment payments are actually made. The total import licensing under special arrangements was of the order of Rs. 192 million.[2] We have assumed that the utilization concentrated entirely in 1954–55 was Rs. 160 million, out of which Rs. 40 million was remitted in the same year (under imports on private account). For the years 1955–56 to 1958–59, negative entries of the following amounts are shown to correct for inclusion of installment payments under imports on private account: 1955–56, Rs. 30 m; 1956–57, Rs. 30 m; 1957–58, Rs. 20 m; and 1958–59, Rs. 20 m. Strictly speaking, the above payments amounted to a repayment of debt and have thus, in our figures, been shown as an offset to foreign loans on the financing side.

The figures on foreign investment covered under non-bank assets relate to cash investment only, i.e., foreign exchange brought in by foreign investors and surrendered to the banks against Pakistan rupees. The statistics relating to imports financed by direct foreign investment are not available. Their

1. The Budget of the Government of Pakistan, 1958–59, Economic Survey and Statistics for 1957 (Karachi: Manager, Government of Pakistan Press, 1958), p. 64.
2. Cf. J. Russell Andrus and Azizali F. Mohammed, The Economy of Pakistan (London: Oxford University Press, 1958), p. 289.

magnitude, however, is not very large.

The suspense item, representing the difference between the current account balance and financing items, is generally due to transitional items or special balancing entries such as book changes in the value of foreign securities and balances.

Appendix Table I.

Monetary Developments in Pakistan

(Fiscal Years: April–March)

(Millions of rupees)

Data for Monetary Analysis	1951–52	1952–53	1953–54	1954–55	1955–56	1956–57	1957–58	1958–59
1. Current Credit Monitization	299	512	457	553	514	1376	1262	1102
2. Internal Monetization (2+5)	301	442	297	289	207	625	548	211
3. Gross Bank lending	339	464	378	427	244	651	634	259
4. Contraction in domestic bank liabilities other than money	-38	-22	-81	-138	-37	-26	-86	-48
5. External Monetization	- 2	70	160	264	307	751	714	891
6. Relative contraction of monetary liquidity (7–8)	- 5	197	- 401	-802	353	322	- 8	533
7. Additional money requirement on the basis of last year's monetary ratio	254	-37	- 23	- 580	971	832	254	679
8. Actual increase in money circulation	259	-234	378	222	618	510	262	146
9. Total monetary financing (1+6), absorbed by:	294	709	56	- 249	867	1698	1254	1635
10. A. Production growth	- 3	145	218	171	-12	289	4	1
11. B. Internal price increases	256	- 182	-239	-751	983	543	250	679
12. C. Excess of imports over exports	35	792	75	315	-188	924	958	959
12a. D. Suspense item and errors omission in balance of payments	6	-46	2	16	84	- 58	42	- 4

Appendix Table II

Monetary Development in Pakistan

(Fiscal Years: April-March)

(In % of Money Supply at the end of the preceding year)

	Data for Monetary Analysis	1951-52	1952-53	1953-54	1954-55	1955-56	1956-57	1957-58	1958-59
1.	Current Credit Monetization	9.1	14.4	13.8	15.0	13.1	30.4	25.0	20.8
2.	Internal Monetization (2+5)	9.1	12.4	9.0	7.8	5.3	13.8	10.9	4.0
3.	Gross bank lending	10.3	13.0	11.4	11.5	6.2	14.4	12.5	4.9
4.	Contraction in domestic bank liabilities other than money	-1.2	-0.6	-2.4	-3.7	-0.9	- 0.6	- 1.7	-0.9
5.	External Monetization	0.0	2.0	4.8	7.2	7.8	16.6	14.1	16.8
6.	Relative contraction of monetary liquidity (7-8)	- 0.2	5.5	-12.1	- 21.7	9.0	7.0	- 0.2	10.0
7.	Additional money requirement on the basis of last year's monetary ratio	7.7	-1.0	- 0.7	- 15.7	24.8	18.3	5.0	12.8
8.	Actual increase in money circulation	7.9	-6.5	11.4	6.0	15.8	11.3	5.2	2.8
9.	Total monetary financing (1+ 6) absorbed by:	8.9	19.9	1.7	- 6.7	22.1	37.4	24.7	30.8
10.	A. Production growth	- 0.1	4.0	6.6	4.6	- 0.3	6.3	0.0	0.0
11.	B. Internal price increases	7.8	-5.1	- 7.2	-20.2	25.1	12.0	4.9	12.8
12.	C. Excess of imports over exports	1.0	22.3	2.2	8.5	- 4.8	20.4	19.0	18.1
12a.	D. Suspense item and errors omission in balance of payments	0.2	- 1.3	0.1	0.4	2.1	-1.3	0.8	- 0.1